

Risk and the inadequacy of science

There are many aspects to risk, both objective and subjective. Difficulties in quantification, compounded by conflicts in value judgements, necessitate increased openness in its assessment, while avoiding oversimplification.

FOR scientists working on phenomena related to risk, 1996 proved to be an unusually challenging year. Climatologists exchanged heavy bombardment with industrial lobby groups, and advisers on the genetic modification of crops and on spongiform encephalopathies found themselves caught up in issues in which important groups (not to mention nations) had much at stake (see pages 6–11). The way forward in such debates, one might think, is to analyse and quantify the risks at the heart of them, and use the results as a basis for prioritizing problems and drawing up solutions. The plea from a British government minister that scientists develop an equivalent of the Richter scale for risks is a reflection of the need for such communicable insight. But the closer one looks at many risks — especially those giving rise to controversy — the more elusive attempts to summarize them seem to become. The minister is doomed to disappointment in his apparent naiveté, but there are constructive ways forward nevertheless.

The quantification of risk would seem to be straightforward enough when applied to a well-understood technology. For example, as was usefully reviewed in the 1992 report of the United Kingdom's Royal Society, *Risk: analysis, perception and management*, engineers can model and test pathways by which a given hazard might arise in industrial plant, and biologists can similarly estimate the impacts of toxins and drugs on the human or animal body.

But judgements have to be made at many stages, even in supposedly objective studies. Extrapolation from animal models to human impacts of drugs carries uncertainties; apparently reasonable assumptions can turn out to be disastrously wrong — witness, for example, the significant underestimate of the impact of Chernobyl fallout on livestock due to faulty extrapolation of mobilities of elements in clays, or the even more notorious misjudgements underlying the explosion of NASA's Challenger shuttle. As was emphasized by the US National Research Council in its recent report *Understanding Risk*, even selecting the appropriate measure of impact needs careful consideration: accidental deaths of employees per ton produced by the US coal industry have been falling in recent years, but deaths per employee have been rising. And where mechanisms are poorly understood and epidemiological data sparse, as is the case in the various types of Creutzfeldt–Jakob disease and the associated prion hypothesis, then it is inevitable that scientific judgement is uncomfortably imprecise. In short, even highly objective analysis of risks is overlaid by judgements that need to be scrutinized.

Then comes the problem of perception, and what the insiders call “multidimensionality”, leading to different responses by people to a perceived risk. A commonly cited diagram in risk studies places specific hazards in a space whose x-axis runs from poorly to well known or understood, and whose y-axis ranges from controllable or voluntary to uncontrollable and involuntary. Almost regardless of relative probabilities, risks are much more likely to be accepted if they lie in the bottom left quadrant (familiar and voluntary), for example sports, than if in the top right — radioactive waste and genetic manipulation. And, of course, official quantifications of risk will be distrusted by the public once experts have been proved wrong (Chernobyl fallout,

BSE, infected blood) or if those responsible for quantifying the risks have strong vested interests (the nuclear industry in waste disposal, Shell with the Brent Spar, Ciba-Geigy with genetically modified corn). As the Challenger disaster illustrated, even the best-intentioned and well-resourced organizations can tend to severely underestimate risks in complex new technologies.

Most people who worry about these things have concluded that openness and participation by non-scientists in the process of risk evaluation is essential. It would be misleading to call this a consensus-generating process. A detriment to some can be a benefit to others — another aspect of risk's multidimensionality. Agreement over how to apply value judgements in evaluating such risks is bound to be elusive — witness debates between developed and developing countries on the value of human lives potentially lost through climate change. But those involved in providing scientific advice have found that including non-experts at the outset sets a broader scientific agenda that gives rise to advice that is all the more useful. The use of peer review and the circulation of preliminary drafts of advice can highlight assumptions that are hidden or poorly accounted for. And the involvement of interest groups ensures not only that all viewpoints have had their say but also that a wider appreciation of the factors underlying the risks and their uncertainties has been gained.

As the NRC's report emphasizes, such openness and participation will not solve all the problems — partly because of differences of principle amongst those involved, and partly because it can be impossible to take into account those whose agenda is directly at odds with the project. A company that feels it has nothing to gain from broad risk estimation and even a lot to lose might choose to boycott the exercise altogether and, instead, hijack the regulatory process that follows it. The communication of risk can also be hijacked, as Greenpeace accomplished in the case of the Brent Spar, making the engineering assessments of risk undertaken by Shell effectively redundant as far as the public were concerned.

All of this shows that attempting to reduce risk to a universal scale is a dangerous notion. On the other hand, understanding of the risks underlying a hazard to public health, or to ecosystems, or to nations, is, after all, an essential part of the political process. There is no alternative to the time and expense of developing such understanding, and communicating it effectively in a way, specific to each case, that allows the media to encapsulate its multidimensional facets. Standards of best practice need to be generated as a goal of regulatory policy. The NRC has provided an excellent step forward in that process, particularly in its report's vivid deployment of examples and case studies.

It is not uncommon to hear scientists (and others) say that, if only people understood the science, their perceptions of risk would be much improved. But that sentiment is misconceived. Lay people can understand the science that they wish to more readily than many scientists seem to credit. What is more critical is that interested scientists and non-scientists alike have examined the components, uncertainties and perceptions of any particular risk. That in turn will focus attention on the relevant science. In that sense, improving the public understanding of risk is a particularly appropriate goal for 1997. □



Reviews of UK university research 'have helped to raise standards'

London. Efforts over several years to concentrate support for research in British universities on those institutions that demonstrate that they are able to use it effectively appear to be bearing fruit, according to the funding councils that distribute government money to them.

The conclusions are based on the fourth research assessment exercise of British universities, the results of which were published in London shortly before Christmas by the Higher Education Funding Council for England. The previous exercises were carried out in 1986, 1989 and 1992, and funding council officials say that the latest round reveals a significant increase in the overall quality of research compared to 1992.

Using a complex and highly codified set of assessment techniques, each department in each university was reviewed by a panel of specialists in the field and given a rating of between 1 and 5, with the third level being split into 3a and 3b, and outstanding departments receiving a 5* mark.

The marks will now be used by the three separate funding councils (covering England, Scotland and Wales respectively), as well as the Department of Education for Northern Ireland, in allocating government funding to institutions for next year. Those given a 1 or 2 rating will not receive any government support for their research (although they can use funds received from elsewhere for this). Those allocated 5* can expect to see continued generous backing.

Perhaps inevitably, not all the marks have been well-received, particularly by departments that feel their research strengths have been underestimated. David Southwood, for example, head of the physics department at Imperial College, London, which was awarded a 5 — but not a 5*, which only went to the departments at Oxford and Cambridge — claimed that insufficient recognition had been given to the breadth and "cohesiveness" of work in his department.

There was deeper anger at the University of Lancaster that their department, which has the highest 'impact factor' of any English physics department as measured by the Institute for Scientific Information in Philadelphia, only received a 3a rating. Peter McClintock, professor of physics, asks: "Why should a department that, averaged over all its staff, is demonstrably England's star performer in international physics have been treated in this shameful way?"

Conversely, the high scores won by some university departments have been looked at slightly sceptically by those at other universi-

ties, who claim that the results may reflect carefully honed submissions rather than quality in depth (universities were allowed to choose which researchers in a particular department to enter).

The final conclusions also gave rise to some public disputes about their interpretation. Both the universities of Oxford and Cambridge claimed to have come top of the list, the first using as an indicator the average score for all the researchers whose work it submitted for review, the second a different score based on all its academic staff.

Overall, however, the level of complaints has been relatively low. "I don't know of any case where a department's score has been more than one grade higher or lower than had been expected," says one biologist.

The results have been a welcome boost to some of the so-called 'new universities', which were called polytechnics until the beginning of the 1990s and are struggling to build reputations in research. Nottingham Trent University — previously Trent Polytechnic — for example, saw 82 per cent of its research-active staff rated in 3a or 3b departments.

"I am delighted with our result," says Richard Joyner, the university's director of research. "It shows that in the new university sector there is a will to develop and manage research effectively."

Brian Fender, former vice-chancellor of the University of Keele, emphasizes that the overall number of university departments

receiving a 3a or 3b rating — each equivalent to a level of 'national excellence' — rose from 589 in 1992 to 716 in 1996, with an increase from 96 to 351 among the former polytechnics alone.

"The real winner in all this appears to be the United Kingdom, which is benefiting from a huge range of high quality work," says Fender. While the best universities have improved their international standing — those receiving grades 5 and 5*, indicating an international level of excellence, rose from 308 to 497 — "many other institutions have performed impressively by identifying and building on their strengths," he says.

Fender also counters claims that some universities have been seeking to increase their scores by selectively recruiting talented individuals at inflated salaries, often referred to as the 'transfer market' by analogy with the activities of professional soccer clubs. He claims that the impact of the transfer market appears to be 'minimal', although anecdotal evidence from many university departments suggests that it remains significant in certain cases.

The assessment results have been welcomed by the lobby group Save British Science. "The exercise shows that we have a substantial asset in our science base," says John Mulvey, the group's secretary. "That is not an excuse for saying that we do not need more money; but it does confirm that, with a bit of extra funding, Britain could become the laboratory of Europe." □

Austria bans gene-modified maize

London. The Austrian government imposed a national ban last week on imports of genetically modified maize from the United States, despite a previous decision by the European Commission (EC) in Brussels to allow the grain to be sold within the European Union (EU).

Two weeks ago, Ritt Bjerregaard, the EC's environment commissioner, accepted the judgement of three of the commission's scientific committees, which concluded that the maize does not pose a threat to human or animal health. But Christa Krammer, Austria's health minister, then announced that Austria would not allow imports of modified maize, claiming to be unconvinced by statements about its safety.

The minister said that the effect of

inserting a marker gene resistant to the antibiotic ampicillin has not been properly evaluated. The maize is also modified to produce a toxin that kills the European corn borer pest, and contains a gene conferring resistance to a herbicide.

The EC decision will satisfy companies, such as Ciba Geigy, which had been prevented from exporting maize to the EU after it had emerged that they intended to sell unsegregated consignments of genetically modified and unmodified grain. Bjerregaard is reported to have promised to revise a directive in which labelling is required only for food that carries specific risks, and Ciba Geigy to have promised to label its bags of modified maize.

Ehsan Masood

Staff pay cost of keeping CERN collider on target

Munich. Officials at the European Laboratory for Particle Physics (CERN) in Geneva, Switzerland, say they will be able to build the Large Hadron Collider by 2005 as planned, despite the governing council's decision shortly before Christmas to accept a substantial budget cut.

This renewed commitment to the agreed timetable makes it almost certain that the United States will contribute towards the US\$2-billion cost of construction. And Japan announced at last month's council meeting that it was prepared to increase its planned contribution from ¥5 billion to ¥8.85 billion (US\$44 million to US\$78 million).

But the price of maintaining the commitment to complete construction in 2005 will be high. Staff will have to accept a pay cut, and there could be a five-year gap in high-energy physics experiments at the laboratory until the collider comes into operation.

CERN council agreed on a series of reductions to the laboratory's planned budget, which would otherwise have been kept at a constant level. The budget will now be reduced by 7.5 per cent below the 1996 level in 1997, 8.5 per cent below for the following three years, and 9.3 per cent after that.

These figures represent a slight compromise on a demand by the German government for an immediate cut of 8.5 per cent, and further cuts to reach the target of 9.3 per cent by 1999 (see *Nature* 382, 387; 1996). Member states promised not to force further cuts on CERN at least until the collider is operational. But the cuts will still require harsher belt-tightening than any European-level laboratory has previously experienced.

The basic salary of staff is being cut by 2.5 per cent this year (see *Nature* 384, 603; 1996). The provision for automatic index-linking of salaries has been removed for 1998 and 1999. And a full review of the salary structure is being carried out to identify further ways to cut costs.

Christopher Llewellyn Smith, CERN's director-general, says that many weaknesses in the current salary structures add to the laboratory's costs. He says benefits for ex-

triate are too generous, and promotion for all staff is too automatic. He is confident that "the majority of staff understand the economic situation in Europe", and will therefore accept the inevitability of cuts in salary and benefits.

One CERN researcher says the laboratory's staff are not expected to "go to the wall" from a 2.5 per cent budget cut. Any potential damage, he says, will come from the terms of the deal, such as the CERN council's decision to dip into a cash reserve intended to compensate for future inflation. "Inflation is low, but is likely to increase. Will council stump up the extra cash when this happens? I have my doubts."

The new budget means that CERN will have no cash available to support plans for additional high-energy particle physics experiments in the period between the switching off in 2000 of the current particle accelerator, the Large Electron-Positron Collider (LEP), and the operation of the Large Hadron Collider in 2005. Smaller high-energy physics programmes such as LEAR (the low energy antiproton ring) and the Omega spectrometer have been recently closed to help pay for the collider.

Llewellyn Smith wants member states to help solve this problem. CERN will have no capital to help, but it can provide old equipment and the infrastructure to support experiments if they can be paid for by national agencies. "Not all countries were in favour of the large reductions to the budget," he points out.

Amid the general gloom, Japan's news was encouraging. The council meeting was told that Japan, a non-member state, was prepared to increase its contribution to the collider which it agreed 18 months ago (see *Nature* 375, 169; 1995), subject to approval by the Japanese parliament, the Diet. According to the president of CERN, Hubert Curien, this provided an important psychological boost, as well as a financial boost. CERN hopes to find room for manoeuvre economically by attracting money from outside the organization.

Alison Abbott

Destruction order for transgenic crop after breach of guidelines

New Delhi. Scientists at the Indian Agricultural Research Institute (IARI) have been told to destroy a transgenic vegetable crop after the Department of Biotechnology (DBT) alleged that the experiment was being carried out without its knowledge and in violation of safety guidelines.

This is the first time the department has taken action under the Environmental Protection Act of 1986, which prohibits unauthorized field trials of genetically manipulated organisms. Under the act, those who violate the department's guidelines on such research can be fined or sent to jail.

The IARI scientists, based in New Delhi, have introduced the toxin gene of *Bacillus thuringiensis* (Bt) imported from Japan into a native variety of brinjal (eggplant or aubergine). The aim is to make the brinjal resistant to insect attack.

But P. K. Ghosh, an adviser to the DBT, says the trial violated guidelines that prohibit open field experiments involving any transgenic crop without a permit. Such permits are issued by the department on the recommendation of a national panel that reviews research proposals for environmental and human safety. Only two permits have so far been issued for transgenic trials and IARI has not received either.

The department is concerned that insects and wind might spread pollen grains from the IARI's experimental plot to neighbouring fields and contaminate the wild relatives of the plant with the foreign gene. The department says that the institute should have conducted experiments in a glasshouse before going to the open field.

R. P. Sharma, head of the research team, says that the field was covered by a giant mosquito net to prevent insect pollinators from entering the trial plot. He says that, since no other brinjals were growing in the vicinity, the question of wind spreading the gene to wild relatives did not arise. But the department has told the scientists to destroy the crop immediately. It admits that it lacks details of experiments at ten other institutes to which IARI has supplied the imported gene. An institute in Simla, for example, which lacks the mandatory biosafety committee, has introduced the gene into potato plants. "We do not know if they are growing the plants in the laboratory or in the field," Ghosh says.

According to Ghosh, one research institute is growing genetically engineered cotton in the open without authorization. The department recently closed down an experiment at an institute in Chandigarh after it discovered that a cholera toxin gene was being handled without adequate safety containment.

K. S. Jayaraman

Russian scientists name 'donor of the year'

Moscow. The Russian Association for the Advancement of Science and Education, a body set up jointly in 1995 by Russian members of the European Academy and the Russian Academy of Sciences, has chosen prominent businessman Boris Berezovsky as its first 'philanthropist of the year' for his support for Russian science.

Berezovsky, currently deputy chairman

of Russia's Security Council with responsibility for negotiations with Chechnya, was formerly chairman and main stockholder of the Logovos company, and is said to be one of the richest businessmen in Russia. Last year he donated US\$1.3 million to a programme that supports travel to conferences, enabling 1,530 Russian scientists to attend international meetings.

Carl Levitin

Debate deferred on proposed database protection treaty

London. The World Intellectual Property Organization (WIPO) in Geneva has postponed until later in the year discussion of a controversial proposed treaty on the intellectual property rights involved in the use of databases.

The treaty was to be discussed at WIPO's conference on Certain Copyright and Neighbouring Rights Questions last month (see *Nature* 383, 653; 1996). But at the end of the meeting delegates merely agreed to recognize the need to "strike a balance" between protecting the interests of both producers and users of databases.

Database compilers want to see agreement on a treaty that will outlaw unfair copying. But users — including many scientists — are concerned that tough regulations may restrict their access to important information. The issue will now be discussed at an extraordinary session of WIPO governing bodies during the next three months. □

EMBL escapes funds squeeze

Munich. Molecular biology has survived the Europe-wide squeeze on funding of international laboratories. At the end of December the council of the Heidelberg-based European Molecular Biology Laboratory unanimously confirmed its commitment to an increase in funding of 3 per cent in real terms next year. It had been feared that Germany or France might try to force a cut to the agreed long-term financial plan. The council also extended the term of office of Fotis Kafatos, the laboratory's director-general, by six years, to his retirement in 2005.

The European Molecular Biology Organization, which operates a programme of research fellowships and conferences, was awarded a 2.6 per cent increase in real terms for next year by its member states. This will allow the organization to increase the number of long-term fellowships it offers by more than 10 per cent. □

New head for DoE research

Washington. The US Department of Energy has appointed Peter Rosen, formerly dean of science at the University of Texas at Arlington, associate director for high energy and nuclear physics. Rosen is a British-born, Oxford-trained particle physicist, who took US citizenship in 1972. He replaces John O'Fallon, who had been in the post on an acting basis. High-energy and nuclear physics are the largest civilian science programmes in the energy department, with a joint annual budget of US\$1 billion. □

Radiation threatens Mars delay

Washington. A manned mission to Mars is likely to have to wait for another quarter century as researchers answer questions on the radiation hazards to humans travelling to the red planet, according to a report from the National Research Council. An earlier launch would require radiation shielding to protect against the highest estimated risk. That could cost NASA US\$10 billion to \$30 billion.

The alternative, says the report, is for the agency to "fund additional research into the biological effects of radiation in deep space". It adds that extended exposure to high energy galactic cosmic rays is likely to result in increased cancer risks, and damage to the brain. "It will take probably more than a decade to answer questions about health risks and the needed protective shielding." □

Novartis heads research table

Basel. Swiss pharmaceuticals giants Ciba Geigy and Sandoz finally merged late last month to become the world's largest life-science company, Novartis, with an annual research and development budget of SFr2.3 billion (US\$1.72 billion). The merger was announced last March, but had to await the approval of the US

Federal Trade Commission, which wanted to ensure that it would not slow research and development or raise the prices of gene therapy products.

As part of the agreement, the commission demanded that Novartis license its patent rights on specific gene therapy technology to Rhône-Poulenc Rorer. The commission had expressed concern that the merger might threaten competition in the US gene therapy market as a whole, and in particular for several specific Novartis gene therapy products: herpes simplex virus thymidine kinase for the treatment of cancer and graft-versus-host disease after bone marrow transplants; factor VIII gene therapy products for the treatment of haemophilia; and interleukins 2, 3 and 6 for cancer chemotherapy. □

New Zealand reshuffles

Sydney. Maurice Williamson has become Minister for Research, Science and Technology as well as for Communications and Information Technology in the new coalition government in New Zealand, the first such government to be elected under the mixed member proportional voting system.

Williamson succeeds Simon Upton who has moved to the environment ministry but retains oversight of the commercially-driven Crown Research Institutes, established as businesses under him when the Department of Scientific and Industrial Research was abolished. Wyatt Creech remains responsible for universities as Minister for Education. He was not tested in the few months he spent in the post before the 12 October election. Two of the nine Crown Research Institutes — Geological and Nuclear Sciences, and Industrial Research — have announced that 60 staff are to be made redundant because of poor profits. □

Herb petrol dismissed as 'fraud'

New Delhi. Ramar Pillai, who found fame with his claims of having produced petrol by mixing water with a secret herb (see *Nature* 383, 112; 1996), has been declared a fraud by India's parliamentary committee on science and technology.

Pillai's dream of an inexpensive solution to the energy crisis was shattered by the report the committee wrote after watching his third experiment at the Indian Institute of Petroleum in Dehra Dun. The report, tabled in parliament, said the experiment was "nothing other" than dissolving camphor in kerosene and passing off the combustible solution as herbal petrol. The herb played no role in producing the end product. Scientists at the institute produced an identical product by repeating the experiment without the herb. □

Time honours AIDS pioneer

Washington. David Ho, the virologist and director of the Aaron Diamond AIDS Research Center in New York, has been named 1996 Man of the Year by *Time* magazine for his contributions to AIDS research (right). The magazine said that Ho's work on therapies that combine several drugs had "raised hope that the virus might some day be eliminated".

Ho established that the AIDS virus, far from lying dormant, reproduces itself in vast numbers during the early stages of infection (see *Nature* 373, 123; 1995), and he pioneered the aggressive use of combination therapies to fight the virus at that stage. Ho, 44, was born in Taiwan and studied physics at the Massachusetts Institute of Technology and California Institute of Technology before entering Harvard Medical School. □



Science advice: a year of dangerous living

From mad cows to 'Gulf War syndrome', 1996 was a year in which politicians faced increasing public demands for action on issues where scientific evidence pointed to a potential danger to human health but remained inconclusive on its precise nature.

Providing advice to government on technology-related risks has long been an important task for the scientific community. Yet inevitable difficulties in providing clear-cut leadership on issues where the interpretation of research results remains uncertain have put strain on the conventional mechanisms for providing scientific advice. So, too, has growing awareness of the role of subjective factors in risk perception.

A series of crises has stimulated governments to develop more sophisticated mechanisms for integrating both scientific advice and public perceptions into their risk-management policies. They have also generated active debate on the relationship between

governments and their science advisers, on the impact of private sponsorship on science advice and on the role of the mass media in moulding public perceptions of risk.

On the following pages, *Nature's* correspondents around the world assess some of the key events and trends around which these debates have taken place in the past year. They range from the growing acceptance of genetic engineering in Germany and the fallout from Japan's food poisoning crisis, to the likely role of US scientists in framing new air pollution regulations.

The contexts — and issues — are very different from one another. But common threads, such as the role of governments in facilitating but not necessarily establishing consensus, and the value of openness and transparency (not forgetting peer review) in preparing science advice, are beginning to emerge; 1997 will show how rapidly they develop. □

How BSE crisis forced Europe out of its complacency

Paris. Science stirred up a political storm last March when Stephen Dorrell, the British health minister, announced that the probable explanation for ten cases of a new variant of Creutzfeldt-Jakob disease (CJD) was that those affected had eaten beef contaminated with the agent causing bovine spongiform encephalopathy (BSE).

Dorrell's announcement to the House of Commons — based on the conclusions of the government's Spongiform Encephalopathy Advisory Committee (SEAC) — brought to an abrupt and ignominious end a decade of reassuring messages from the government. These had relied on the premise that there was no scientific evidence that BSE could pass to humans and that, even if it could, adequate measures were in place to prevent the most infective parts of cattle entering the human food chain.

The UK government's U-turn plunged Europe into one of its biggest economic and political crises since the Second World War. The European Union (EU) imposed a ban on all UK exports of beef (see *Nature* 381, 354; 353; 1996), and consumers panicked at the sudden prospect of a CJD epidemic. Beef consumption across the EU in 1996 was 11 per cent down on 1995. By the end of this year, the BSE crisis will have cost the EU ECU3.5 billion (US\$2.8 billion) in subsidies to the beef industry.

The impact of the crisis on the public's esteem for scientific advice has been no less dramatic. The main message that the public appears to have retained is that, despite warning signals, structural weaknesses in the operation of the institutional framework meant that a dangerous complacency permeated the government, the scientific and medical communities and the media.

Investigations by both the press and the European Parliament have unearthed a

catalogue of errors by UK and EU authorities — in particular the failure adequately to enforce either abattoir controls or a 1988 ban on the feeding of ruminant protein to ruminants (see *Nature* 381, 544; 1996).

All this has led to a cacophony of demands for change in the way in which Britain and the EU organize the gathering of scientific advice (see *Nature* 384, 201 & 8; 1996). Several common threads seem to be emerging. One is the desire to introduce a

to take over food regulation from the agriculture ministry. A similar separation of powers within the European Commission is likely to feature among the recommendations of a BSE inquiry set up by the European Parliament, which is scheduled to release its final report next month.

One proposal is expected to be to remove animal health and food safety from the agriculture and industry directorates, combine them with health — currently part of the directorate for employment and social affairs — and transfer these responsibilities to the directorate for consumer policy (see *Nature* 384, 8; 1996).

Pressure is also mounting in Britain and elsewhere for reform of the system of government scientific advisory committees. The UK Consumers' Association and Charter 88 (a body campaigning for constitutional reform) have demanded greater accountability by such committees (see *Nature* 384, 201; 1996).

One unpublished account of the operation of the commission's scientific veterinary committee will give critics further cause for concern. A 1990 memorandum written by one commission official, Gérard Castille, records how Fernando Mansito, deputy director general of the agriculture directorate, reprimanded the committee at a meeting on BSE in June 1990 for "only discussing free circulation [of products] and economic aspects", although its remit was to assess the risks to consumers. "On two occasions, he was obliged to remind the committee of its mandate," says the memorandum.

Franz Fischler, the European agriculture commissioner, has himself publicly stated that a reform of the commission's scientific advisory system is needed to re-establish its credibility and impartiality in the eyes of the public.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Eye of the storm: farmers protest to the UK parliament (above) while Britain's chief vet, Keith Meldrum (left), faced questions in Brussels.

clearer institutional separation between public health and consumer protection on the one hand, and agricultural and industrial interests on the other.

Decision-making on BSE in the United Kingdom has been concentrated within the Ministry of Agriculture, Fisheries and Food, and at the European level within the European Commission's directorate for agriculture. Critics argue that this arrangement has allowed agricultural and economic interests to prevail over those of public health (see *Nature* 380, 273; 1996).

In the United Kingdom, both the opposition Labour party and consumer associations have called for an independent body

Fischler has suggested that management of scientific committees within the commission might be better transferred to independent agencies, as in the United States, where the government often asks the National Academy of Sciences and the Institute of Medicine to report on controversial topics.

He also argued last month that new mechanisms are needed to prevent the narrow interests of individual member states from prevailing over science and the collective interest. Fischler has floated the idea of creating an independent food agency. But it is far from clear that this is a practical proposition, as such an administrative structure would have few powers under the current EU statutes.

To the same end, the parliament would like decisions on public health to be managed at a higher political level. Parliament officials argue that the nomination of a single commissioner for all aspects of public health would mean that issues such as BSE would require high profile negotiations with the agriculture and industry commissioners.

Another demand that has been made strongly by groups in the United Kingdom and in the parliament is the need to broaden the expertise of scientific committees. The predominance of veterinarians on the UK and the commission's BSE committees resulted in blinkered vision, claims one official at the parliament.

The BSE crisis has revealed the need to include on such committees researchers from other disciplines, such as molecular biology, argue critics. Recognizing this, the commission last summer established a multidisciplinary committee on BSE which may serve as a model for the gathering of scientific views on other issues in future (see *Nature* **381**, 724; 1996).

But, whatever reforms are introduced in 1997, the commission's scientific advisory and regulatory systems risk being dogged for some time to come by the conflicting interests of member states. The need to reach compromises may prevail over purely scientific considerations.

As if to prove this point Fischler has over the past few months championed the conclusion of the commission's scientific advisers on the need to ban sheep offals, on the grounds that sheep may be harbouring BSE.

The response of the member states was fairly predictable. While Britain and France favour a ban, it has been opposed by other members, including Germany, which argue that it is unnecessary given that they have ostensibly no cases of BSE. Despite the lessons of BSE, it seems that science will continue to have an uphill struggle in 1997 in competing with economic and political considerations.

Declan Butler

All *Nature* news stories — including hard links to referenced items in previous issues — are accessible without charge on the World Wide Web at www.nature.com or www.america.nature.com.

Climate panel forecasts way ahead

London. How can science advice on controversial issues effectively feed into the policy-making process? One pioneer has been the Intergovernmental Panel on Climate Change (IPCC). Despite being dogged by controversy which escalated considerably during 1996, the IPCC appears set to become a model for other areas of risk-based policy-making.

The IPCC was set up in 1988 by the World Meteorological Organization and the United Nations Environment Programme. It involves many of the world's leading climate researchers who prepare comprehensive five-yearly reports on the state of the world's climate.

Its reports have a direct influence on global climate policy. The latest 3,600-page report ended a long-running debate on whether human activities, such as burning fossil fuels, were responsible for the rise in average world temperatures (see *Nature* **378**, 524; 1995).

The report was endorsed with varying degrees of enthusiasm by around 100 countries which have signed the climate convention. It was also backed by the United States, the world's largest emitter of carbon dioxide, opening the door for a recognition of the need for policies to reduce carbon dioxide emissions (see *Nature* **382**, 287; 1996).

The IPCC's achievement is partly due to its unique structure, in which scientists representing both independent research institutes and national governments are given the chance to contribute. Sir John Houghton, chairman of Britain's Royal Commission on Environmental Pollution and co-chair of the IPCC's science working group, says direct government involvement in the process not only provides the reports with political clout but also gives scientists direct access to policy-makers.

The IPCC is divided into three working groups. The first deals with the science, the second with impacts, and the third with the economics of climate change. Each group is headed by two chairmen, drawn from a developed and a developing country. Working groups are further sub-divided into chapters, comprising many 'contributing authors' and headed by a team of expert 'lead authors' drawn from independent research institutes.

The work of each chapter is reviewed by a different group of independent experts, government scientists, and non-governmental organizations before being finalized by the lead authors. Government scientists are then responsible for line-by-line approval of a 'summary for policy-makers' of the entire IPCC report.

Close attention is paid to this summary,

as it is often the only document read by policy-makers. The lead authors also write their own summary. Efforts are made to ensure all three documents agree. But success is not guaranteed.

Government scientists often come with prior agendas reflecting national interests. The oil states, for example, remain sceptical of measures for reducing carbon dioxide, which will hit their export income. The small island states, on the other hand, are lobbying for tough measures because their existence would be threatened if sea levels were to rise.

Because climate science is complicated and largely determined by computer models of future climate, tensions



inevitably emerge over interpretation and emphasis. Developing countries, for example, remain opposed to the methodology used in one chapter to calculate projected loss of life and property from climate change, in which losses in the developed world were valued at a higher rate than for less developed countries. Last year developing countries' scientists refused to endorse this principle when drafting the IPCC document's summary for policy-makers (see *Nature* **378**, 429; 1996).

In another example, Saudi Arabia and Kuwait lobbied to water down conclusions on the human influence on global climate. They eventually relented after being marginalized by the rest of the drafting committee. But their campaign was taken up by a prominent US fossil-fuel lobby, which protested when it learnt that parts of the main report had been altered after peer review to clarify ambiguities that had arisen during the writing of the policy-makers' summary (see *Nature* **381**, 546; 1996).

Despite the setbacks, Houghton believes the IPCC has proved a success. "The science must never be compromised. But scientists need to be patient with politicians."

Ehsan Masood

UK policy learns about risk the hard way

London. When Ian Taylor, Britain's science minister, addressed the British Association for the Advancement of Science last September, the subject he chose was the way public concern about technological risks is increasing scepticism towards science.

In his address, Taylor advocated the use of a 'Richter scale' of risks to communicate the relative severity of different types of events. His proposal reflected frustration in government circles over the public's apparent refusal to adopt a rational approach based on a scientific assessment of hazards.

But, in passing references to other elements — such as the 'dread factor' — Taylor also indicated a growing, if sometimes reluctant, acceptance of what sociologists have been telling governments for some time: that an effective approach to the public perception of risks must take into account the importance of less 'rational' factors.

Britain has learnt its lessons the hard way. At the start of last year, the government was still smarting at the failure of the Shell oil company to stand firm in the face of environmentalist protests over the dumping of the Brent Spar disused oil storage platform in the North Sea (see *Nature* 378, 376; 1995).

But the biggest blow came over bovine spongiform encephalopathy (BSE), and in particular the announcement in March by Stephen Dorrell, the health secretary, that government science advisers were no longer

ruling out a possible connection with human Creutzfeldt-Jakob disease (see page 6).

The announcement triggered a flood of ridicule of previous statements about the safety of beef for which the government paid a heavy political price. The economic costs were also enormous.

The crisis has pulled the rug from under the feet of those who argue that the views of scientific (and economic) experts are a sufficient basis for an effective approach to the political management of risk.

It has been the Treasury — which has had to pick up the bill for previous failures — that has been the prime motivating force behind efforts to develop a more effective approach. Pressure from the Treasury has led to the creation of an Interdepartmental Liaison Group on Risk Assessment, run by the Health and Safety Executive, which is aiming to develop standard government-wide techniques for handling risk.

Some individual advisory groups have already departed significantly from the previous approach of drawing primarily on 'hard' science, and are seeking to broaden the base on which their advice is established. An example is the Advisory Committee on Novel Foods and Processes, which warned the government about the potential dangers of an antibiotic marker left in genetically engineered maize grown in the United States (see *Nature* 383, 564; 1996). The com-

mittee's chairman, Derek Burke, a microbiologist and former vice-chancellor of the University of East Anglia, says the advice of his group reflected the views of some members that any increase in antibiotic resistance — however small — should be opposed.

"Experiences such as that over irradiated foods, which were approved but were rejected by consumers, have changed the way we think about our role," says Burke. "I am much more sensitive to consumer issues than I was five or six years ago."

Shell has also learnt from the lessons of the past. Since last year's retreat on Brent Spar, the company has transformed its approach and is now engaged in an extensive dialogue with environmentalist groups and others on platform disposal.

Meanwhile, some government officials are pointing to the consensus-generating strategy of the Intergovernmental Panel on Climate Change (see page 7) as a guide to the way in which contentious scientific issues should be handled.

Not everyone has welcomed the new approach. Publication of the report of a panel of advisers on Brent Spar was apparently delayed because Tim Eggar, the energy minister, was uncomfortable about its failure to condemn environmentalists' arguments outright (see *Nature* 381, 99; 1996). Right-wing think-tanks such as the Institute for Economic Affairs continue to

German resistance to genetic engineering diminishes

Munich. Genetic engineering remained high on the public agenda in Germany last year. With the Nazi abuse of genetics still fixed firmly in the German conscience, a conference marking the 50th anniversary of the Nuremberg war crimes trials attracted wide publicity and helped to keep the debate going (see *Nature* 384, 5; 1996).

At the same time, however, 1996 was a year in which opposition to genetic engineering appears to have lost some ground. Although its agricultural uses remained under suspicion, medical uses found much wider acceptance. Even the Green party admitted that, under certain circumstances,

genetic engineering could be beneficial.

An opinion poll published last summer by the Institute for Public Opinion Research in Allensbach confirmed this growing public acceptance. Only 29 per cent of about 1,000 individuals questioned said Germany should reject engineering completely because the risks were too high, compared with 40 per cent in a similar poll eight years previously.

Gerd Hobom, president of the Zentrale Kommission für Biologische Sicherheit (ZKBS), the federal office that controls licensing of all genetic experimentation, says that one important factor in boosting the acceptance of genetic engineering has been the BioRegio competition launched by Jürgen Rüttgers, the federal research minister (see *Nature* 384, 298; 1996). In preparing their entries for this competition, local governments put aside their doubts and backed collaboration between academics and local industries to develop regional biotechnology projects.

Such cooperation even took place

in the state of Hessen, where plans for Hoechst's production plant for genetically engineered insulin were blocked by the environment ministry for nearly a decade (see *Nature* 360, 402; 1992). Dieter Brauer, a spokesman for Hoechst, the chemicals and pharmaceutical company, says that the notorious obstacles facing such companies in Germany have virtually disappeared. One reason, he says, has been changes to the 1990 Gene Law that reduce approval times for new projects and facilities (see *Nature* 367, 210; 1994).

The most important factor, says Brauer, is a change in attitude — largely due to the economic recession and a relatively high level of unemployment. Nevertheless, resistance to the agricultural uses of genetic engineering remains firmly entrenched. Last year it was revealed that almost all field trials of genetically-manipulated crops in Germany had been either totally or partially destroyed by protesters (see *Nature* 380, 94; 1996). And Germans remain stubbornly opposed to genetically modified foods.

Unlike countries such as the United Kingdom, Germany has no formal mechanism for integrating public perception of risks into political decision-making. Further-

Red menace? Protests against genetic engineering (here in Berlin) continued to reflect public concern about its use in agriculture.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Blood scandal and *E. coli* raise questions in Japan

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS
UNAVAILABLE
FOR COPYRIGHT
REASONS

All at sea? Brent Spar pointed up differences over policy based on 'hard' science alone.

argue that government policy on topics such as climate change should be based only on proven scientific 'fact'.

But others, such as Brian Wynne, an expert on risk at the University of Lancaster, argue that attempts to require standard risk management techniques across all sectors of government can ignore important differences in the social and political contexts in which risks arise. If the BSE crisis and other events of 1996 have revealed how not to do things, 1997 will be watched closely to see how widely the lessons have been learnt.

David Dickson

more, problems tend to be addressed in a traditional consensus-seeking manner, whereby agreement is sought through discussions and compromise. Attempts to use this approach with genetic engineering have met little success.

One example is an exercise in 'participatory risk assessment' on the genetic engineering of plants carried out from 1991 to 1993 and involving 50 individuals from environmental groups, industry, regulatory authorities and the scientific community. The meetings were organized by Wolfgang van den Daele, a sociologist at Berlin's Science Centre for Social Research.

But groups critical of genetic engineering withdrew before the last meeting because they were unhappy with the direction of the discussions, which meant that no consensus conclusions could be agreed. The talks had been leading to the conclusion that the central issue of the controversy was the political one of democratic control of new technologies rather than real risks. Critics did not want to see their objections on safety grounds being reduced to a political goal.

The failure of the consensus approach has put strain on attempts to bring public participation into risk-related decision-

Tokyo. Scientists are generally held in high regard in Japan, and their views tend to be accepted uncritically by the public. But that faith has been severely strained by the events of the past year.

Two events in particular have eroded public trust in scientists as arbiters of health risks — the scandal over the contamination by HIV of blood products used by haemophiliacs, and the food poisoning of thousands of schoolchildren by a highly toxic strain of the bacterium *Escherichia coli*.

The year opened with the Minister of Health and Welfare, Naoto Kan, acknowledging the failings of his ministry in not halting the spread of HIV to haemophiliacs (see *Nature* 379, 663; 1996). An investigation he initiated confirmed what had been suspected for years: that, despite discussing emergency action, the ministry had failed to react to warning signs of HIV contamination of blood coagulants in 1983.

Public attention became focused on the key role that Takeshi Abe, formerly vice-president of Teikyo University, had played in advising the government to continue the use of blood products that had not been heat treated to kill viruses. Also under public scrutiny was Abe's continued administration of such products to his patients well into 1985. Both issues led to

making. Indeed, at a meeting in Bavaria early last year organized by Dechema, a group that supports the chemical industries, many supporters of genetic engineering argued that, with the 1990 regulations in place, they no longer needed to take account of the views of critics.

In 1993, the Bundestag, the German parliament, asked its technological assessment group, TAB (Büro für Technikfolgen-Abschätzung), to commission a full risk assessment of genetic technologies. TAB officials say that they are surprised that parliament has not called for a similar study on genetically-engineered foods, given the fierce level of public debate on the topic. Last month, on its own initiative, it therefore commissioned a small study, coordinated by Gert Spelberg of the consumers' organization, Verbraucher Initiative.

Spelberg says that, perhaps as a result of van den Daele's exercise, the consumers' voice is now being heard. He points out that Germany led the controversial move to require detailed labelling of foods containing genetically modified products in European Union legislation that will come into effect this year (see *Nature* 384, 301; 1996).

Alison Abbott

his arrest on charges of wilful negligence resulting in death (see *Nature* 383, 6; 1996).

More bad light was thrown on medical science by the role of other scientists on the ministry's advisory committee. For example, it was reported that Yuichi Shiokawa, who headed the AIDS advisory committee after Abe, had rejected recognition of one of Abe's patients as Japan's first haemophiliac AIDS victim in 1983. Two years later, he proposed a homosexual patient from his own hospital as Japan's first AIDS patient despite minimal evidence.

Exposure of the ways in which such decisions were made has led to an increase in the public's distrust of science. But the public still sees Abe and Shiokawa as exceptions. In contrast, others used the activities of these two to argue that there are serious defects in Japan's whole medical system.

The public row over the contaminated blood scandal was closely followed by the food poisoning of thousands of children by the O157 strain of *E. coli*. In this case, the public turned with concern to the government, scientists and the medical world for guidance and solutions. But no clear-cut answers emerged.

Preliminary DNA analysis suggested many possible sources of infection. But the public demanded more precise explanation. And Kan came under pressure to provide one. His ministry — using only circumstantial evidence — announced that radish sprouts (*kaiware daikon*) in school lunches were the likely source of *E. coli* contamination (see *Nature* 382, 567; 1996).

The announcement brought radish farmers throughout Japan close to the brink of financial ruin. But it was never established that the radish sprouts were indeed a source of the mass infection. And Kan was subsequently forced to eat radish sprouts in public to restore confidence.

To many, the blood scandal and the *E. coli* food poisoning have underlined a basic lack of understanding among the Japanese public of the limitations of science. Some feel that the experience of the Kobe earthquake disaster of 1995 — which the experts failed to predict — should have altered this attitude. But the Japanese still look to science for black-and-white answers.

Some popular politicians such as Kan are making moves towards greater public involvement in decision-making on issues such as health and nuclear power.

But many Japanese scientists feel that, as long as public awareness and understanding of science remains limited, such moves are unlikely on their own to lead to better management of science-related risks.

David Swinbanks

Voice of US science struggles to be heard

Washington. The United States takes justifiable pride in having one of the world's most comprehensive and open systems of scientific advice. The National Academy of Sciences and its associated bodies have 1,500 staff more or less dedicated to the purpose, as well as a plethora of panels of experts. These ensure that the voice of science is heard in the corridors of power.

But over the past year this voice has sometimes had trouble getting itself heard. Indeed, by the time in 1995 that the previous Congress abolished its Office of Technology Assessment — which provided technical advice to members of all political stripes — it was already clear that the stock of technical objectivity was beginning to fall in some political quarters.

The tendency has arisen for interested parties in political disputes on technical issues to ask not “what do the scientists say?” but “who are they?”, “what do they want?” and — perhaps most revealingly — “who pays their salaries?”

Prominent scientists who participate in the activities of the Intergovernmental Panel on Climate Change, for example, are government-funded. Republicans in Congress claim that they therefore must be seeking additional government support for their

research (see *Nature* 382, 195; 1996).

The chair of an external scientific advisory committee at the Environmental Protection Agency (EPA) works for General Motors. So, environmental groups contend, it is not surprising that he is obstructing tougher regulation of air pollution (see *Nature* 380, 11; 1996).

Add to this a general growth in cynicism towards the federal government among the US public, and the stage is set for a slump in the level of esteem in which ‘expert’ scientific advice is held. One lurking iceberg that could inflict such damage on its own is ‘Gulf War syndrome’, the mysterious ailment allegedly affecting US veterans of the 1991 conflict.

The Department of Defense has consistently denied that a conclusive pattern of illness exists among the veterans, and numerous eminent scientific panels have supported this view. But the public, and Congress, have persistently mistrusted the department's position.

A major study for the Defense Science Board in 1994, carried out by a panel that was chaired by Joshua Lederberg, the Nobel prizewinning geneticist, dismissed chemical reagents as a possible cause of the illness. A panel at the National Institutes

of Health reached similar conclusions (see *Nature* 369, 8; 1994).

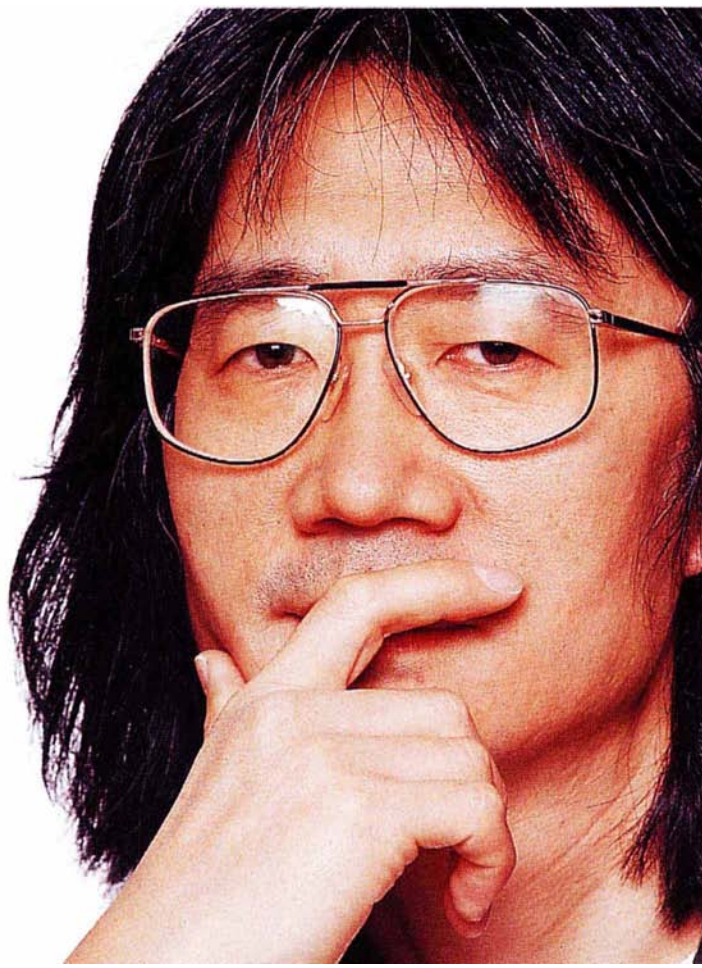
But Lederberg now says that his panel was kept in the dark by the Pentagon about critical events during the war. “We didn’t get all the information, and I don’t know where it was,” he told the *New York Times* last month. “If I had to do it all over again, I would have spent more time and effort digging out the details.” Lederberg has not said that the study was wrong, but his comments are likely to undermine its conclusions. The National Academy of Sciences has now been asked to consider the issue again.

If veterans’ advocates are proved correct and Gulf War syndrome is indeed linked to chemical agents encountered during the war, public confidence in the objectivity of government-appointed scientific panels will inevitably suffer. And, even without such a public humiliation, the role of such panels is likely to continue to come under strain.

An impending battle over new clean air rules proposed by the EPA, for example, seems destined to highlight the weaknesses of the advisory system, rather than its strengths.

The EPA’s Clean Air Scientific Advisory Committee (CASAC) has criticized the proposed regulations as being too harsh.

To **save an hour**
each time you
purify histidine-
tagged proteins,
put it on the
tip of a syringe



But some environmentalists dismiss CASAC as a mouthpiece of industry, and the agency appears to be determined to implement the regulations.

The fate of the proposed regulations will, as a result, depend not on the opinion of CASAC or of other experts, but rather on the outcome of yet another hard-fought battle waged between environmental and industrial lobbyists.

The terms under which these battles are fought have recently been changed by Congress, although not as radically as the Republican majority would have liked. When it passed a risk assessment and cost-benefit analysis act last year, the House of Representatives wanted every proposed regulation to be subject to an exacting process to prove its economic worth. But the Clinton administration successfully blocked the law in the Senate. It argued that, although quantitative risk assessment is a valuable tool, the proposed rule would conflict with the duty of agencies such as the EPA to put public health first and think about costs later.

Clinton appears to have won public backing for his position, and the risk assessment legislation is unlikely to re-emerge this year. But a separate law to protect small businesses, which did pass last year, will give Congress greater opportunity to scrutinize

new regulations from government agencies.

Meanwhile, Republicans say they will continue to attempt to inject the principles of the risk assessment law into other environmental legislation. "We want to get the assumptions made by the regulating agencies out on the table," says one congressional staff member.

Such openness is being demanded by each party in the increasingly heated debate

closer scrutiny by both groups. Last year, the academy proposed a new approach to risk assessment, involving all participants in the process from the outset (see *Nature* 381, 638; 1996). The model met with some scepticism: one industrialist predicted that it would cause science to "bring an even more confusing dog's breakfast to the table" than it does now.

Several government agencies have already started to implement the academy model, or will do so this year. The result will doubtless be more meetings, more consultation, more public argument, more law suits, and more shameless lobbying in Washington for self-serving positions. Critics believe that nobody's interest will be served by all this heat.

But recent history suggests that all the heat will ultimately succeed in shedding some light on the complex risks the federal government must manage. On a wide range of risks from air and water pollution to nuclear and food safety, the American public has tended to enjoy better protection — and quicker application of relevant science and technology — than have the people of Europe or Japan. As the United States' already transparent processes are opened up further, this is likely to remain the case.

Colin Macilwain

IMAGE UNAVAILABLE FOR COPYRIGHT REASONS

Revvig up: does clean air policy depend on lobbying or science?

in the United States on how risks should be assessed. Put bluntly, environmentalists want the process open to public scrutiny — which, in many cases, means scrutiny by environmentalists who are the interested public. Industrialists want to ensure scrutiny by scientists, partly because they employ many of the scientists best qualified for the task.

The government is likely to respond by opening the risk assessment process to even

Environmental Images/C. Jones

HisTrap kit: saving an hour couldn't be simpler

Would you like to save one hour in histidine-tagged protein purification?
To load, purify and elute in 15 minutes? Now it's possible—time after time.

The HisTrap™ Kit presents you with a solution so simple it works with a syringe. Inside the kit is all you'll need for up to ten affinity purification runs: pre-packed cartridges, ready-made stock solutions, a protocol that's easy to follow, and, of course, a syringe.

You'll get high purity recombinant proteins as the kit lets you use the well-known binding of poly(histidine) sequences to immobilized nickel. What's more, it saves you the effort of weighing out buffer salts.

Molecular biologists have been purifying recombinant proteins with Pharmacia Biotech media from the very beginning. Call us at 1 (800) 526 3593 from the USA, +81 (0)3 3492 6949 from Japan or +46 (0)18 16 50 11 from Europe and the rest of the world (or meet us on the Internet at <http://www.biotech.pharmacia.se>). Ask about the kit that saves an hour.

**Pharmacia
Biotech**
Uppsala, Sweden. (And the rest of the world)

A column from
the HisTrap™ Kit,
magnified 280%.



22 96, MAMC

"Purity, simplicity and time savings: a product that gives you all these, lets you take a break from the lab for a bagel, coffee and...hmm, maybe a haircut?" Kyungsun Suh, Ph.D., at the

Unequal struggle in South Africa

Postdoc woes

SIR — I should like to comment on your leading article and on the opening story in the Briefing by Michael Cherry on South African science (*Nature* **384**, 1 & 11–15; 1996). Your leading article rightly recognizes the need for reforms to “redress the idiosyncrasies” of the previous government and that the published policies look good on paper. In the second article, Cherry points out that the published policies have raised the hopes of many South African scientists, yet are slow in taking effect and are likely to be constrained in their effectiveness. Both articles give a good representation of the state of South African science and neither is shy of direct criticism.

Cherry draws attention to the fact that achieving the anticipated “renaissance”, which many South African scientists, in his opinion, perceive to be “just around the corner”, is likely to be constrained by political and financial factors. I agree, but I should like to expand on the political constraints by drawing attention to an overriding policy not mentioned in the articles, namely “affirmative action”.

Unlike Cherry, who I gather is an established academic at one of the five ‘major’ universities in South Africa (which are all historically white universities, of course), I am a PhD student. I also happen to be at an historically black university, where I have been temporarily employed as a researcher for 6.5 years. As I am nearing completion of my PhD, I am seeking to establish myself as an academic so as to contribute to the development of science in the new South Africa. It is in this position that I have realized the full force of affirmative action, because I am white.

Until recently, I had the optimism and high hopes that Cherry mentions, but not any more. Whereas black graduates have access to local and foreign funds, and opportunities to study abroad and get jobs, the same does not apply to their white compatriots. It is only established white South African scientists who can afford to be optimistic. Yet I hear established, world-class South African scientists voicing their concern more and more often about the future for (white) graduates in the country.

I do not begrudge young black South African scientists the long overdue opportunities that are now available, but what happened to ‘equal opportunity’? Is the ruthless application of African nationalist ideology, under the guise of affirmative action, and at the expense of local expertise, really the way to “redress the idiosyncrasies” of the previous government?

It is true that reforms are necessary and that the published policies raise the levels of hope, but it is the anticipated non-implementation and ineffectiveness of the pro-

posed changes and, rather, the implementation of unpublished policies, that concerns me. The last sentence of your leading article, Sir, starts with a very big “If”!

Peter F. Scogings
Agricultural & Rural Development
Research Institute,
University of Fort Hare,
Alice ZA-5700, South Africa
e-mail: scogings@ardri.uth.ac.za

The moral ape

SIR — Frans de Waal’s argument for emergent unselfishness in the great apes and humans is a welcome challenge to the belief that, because genes are selfish, all higher levels of biological activity are fundamentally selfish as well (*Nature* **383**, 785; 1996). I find no conflict in accepting gene-level selfishness and neocortical-level unselfishness in complex brains.

The handful of experts in chimpanzee and bonobo social behaviour (indeed, they can be counted on one hand) all emphasize the ways in which the social behaviour of these primates shares features of human behaviour: apparent concern for community well-being, acts of caring and consolation toward non-kin as well as kin, reconciliation after fights, reciprocal sharing of resources (and retaliation for non-reciprocity) and an apparent social ‘conscience’.

Because of the close genetic relationship of chimpanzees, bonobos and humans, as revealed by DNA–DNA hybridization, gene sequencing and morphology, we should seriously consider the behaviour of these close relatives as being evolutionarily relevant to ours, as de Waal urges (in *Good Natured: The Origins of Right and Wrong in Humans and Other Animals*; Harvard University Press, 1996). Instead of merely being appended atop a selfish core, ‘merely’ behaviour may have become deeply integrated by natural selection into our psychological nature. In my experience as a mental health professional and as a human being, I have observed profoundly unselfish feelings and behaviour which appear as genuine as their selfish counterparts.

Gregory G. Dimijian
University of Texas Southwestern Medical
Center at Dallas,
5323 Harry Hines Boulevard,
Dallas, Texas 75235-9070, USA
e-mail: dimijian@connect.net

corres@nature.com

Letters submitted for Correspondence may be e-mailed to **corres@nature.com**. Do not send items intended for any other section by e-mail unless requested to do so.

SIR — I read with interest the correspondence (*Nature* **384**, 103; 1996) about the difficulties experienced by today’s young researchers. As such a person, I believe postdoctoral experience is extremely valuable to the development of one’s research career, particularly if undertaken at foreign institutions where the differences in culture and methods are the most striking and can be of the most benefit.

But asking someone to make such an undertaking today is I believe, unfair. Most postdoctoral scientists receive wages above the local average, but when moving to a different country this may be far from sufficient. Some institutions still do not include travel or relocation expenses for new staff members. Foreign governments usually prohibit accompanying partners of temporary workers to seek paid employment. As most families need more than one breadwinner, the family income is effectively halved.

Not all countries provide for superannuation or retirement annuity, which is becoming increasingly important because of the down-sizing of social security. Having retirement funds, when available, in different countries is a disadvantage compared to someone with funds in only one country. And moving funds to another country incurs penalties for early withdrawal.

All these problems are exacerbated when young researchers are essentially forced to undertake more than one postdoctoral term because of the relative scarcity of more senior and stable positions. With so many researchers in the world, most of whom have jumped through the same hoops for many years, things might have been expected to have changed for the better by now. Administrators should realize that work and family conditions are different from 20 or so years ago.

We come ultimately to the question of whether obtaining a PhD lives up to our expectations. A PhD is always valuable, but it involves a special level of commitment and love for a particular subject. Most graduates do not complete a PhD so as to become World-Wide Web programmers.

Career and family responsibilities may always have clashed, but young researchers are now asking themselves too often whether to cut their losses and run. The way science is done today has changed and institutions should respond accordingly, for their own sakes and for the sakes of the researchers who support them.

Emanuel Vassiliadis
Instituto de Astrofísica de Canarias (IAC),
c/- Vía Lactea S/N,
E-38200 La Laguna, Tenerife,
Canary Islands, Spain
email: vass@iac.es

Eighteen-ninety-seven and all that

J. L. Heilbron and W. F. Bynum

Anniversaries to celebrate this new year range from the invention of the flute and the sewing machine to the discovery of the mountains of the Moon, some small killers, and an even tinier elementary particle.

THE anniversary of the year is the discovery in 1897 of the first elementary particle, of the bond between physics and chemistry, of the basis of computing, of that faithful and tireless servant of mankind — the electron.

We have arranged our anniversaries as in previous columns in this series, beginning with centennials, sesquicentennials and so on, back to 1247. We end with a list of memorable discoveries made in our century 50, or 50 ± 25 , years ago. Also, as usual, ϵ signifies 'centennial'. We use \$ to indicate a Nobel prize.

1897: Elementary advance

Joseph John Thomson (\$1906) identified what is now called the electron 100 years ago. The usual account of his success describes how he subjected cathode rays to electric and magnetic fields adjusted to give no net deflection.

The equation of the balance of forces yielded a value for the charge-to-mass ratio, e/m , of the particles that, according to most English physicists of the time, made up the cathode rays. Because e/m came out to be 1,000 times the value of E/M , the charge-to-mass ratio of the hydrogen ion as determined by electrolysis, Thomson had the options of $e > E$ or $m < M$ or both. He chose $m = M/1,000$ so as to get a corpuscle small enough to behave like cathode rays. He conjectured that these corpuscles made up not only cathode rays but also all matter.

The conjecture evidently outran the facts. Nonetheless it soon carried conviction. That was not because Thomson had found a way, where no one had before, to exert an electrostatic force on the cathode-ray particles. Rather, it was because a set of investigations — some of them independent of him and the Cavendish Laboratory at Cambridge that he directed — fitted his extrapolation perfectly.

These were the findings that the charge-to-mass ratios of the particles liberated from metals by ultraviolet light, of the ions supposedly responsible for the emission of spectral lines in the Zeeman effect, and of the easily absorbed, or beta, fraction of the rays from radioactive substances, all equalled approximately the e/m of the corpuscle; and that the charges carried by negative gaseous ions, which Thomson supposed to consist of normal atoms bearing electrons, were small multiples of e . By 1900, most physicists had come to agree with Thomson and to recognize that his dis-

Particle convection: Thomson with a cathode-ray tube.

covery had transformed their science.

Among those who had a piece of the action in 1897 were Thomson's student Ernest Rutherford, who that year distinguished between alpha and beta rays; Marie Curie, who identified the uranium atom as the source of "uranium rays"; and Walther Kaufmann, who also obtained a value for e/m for the cathode rays but did not know what to make of it. The corpuscle became the electron when everyone agreed that it carried the unit charge, which had been dubbed the electron in 1891, about the time Thomson took up the study of cathode rays. The ease of exploiting the electron was illustrated in the year of its discovery by Karl Ferdinand Braun (\$1909), who designed a forerunner of the oscilloscope, which, in a modern form, is the picture tube of television sets.

Until 1897, most knowledgeable people believed that the products of organic transformation, such as alcohol, could be made only by a living cell. Then Eduard Buchner (died 1917, 0.8 ϵ) caused a cell-free extract of yeast to act on sugar. The result — that alcohol could be made without the intervention of life — was a great spur to biochemistry and the motivation for his Nobel prize (\$1907, 0.9 ϵ). The discovery that there is no *vie in eau de vie* reminds us that C. W. Post's grapenuts, which contain neither grapes nor nuts, first appeared on breakfast tables in 1897.

Further to the chemical essence of things, Emil Hermann Fischer (\$1902) proposed the structure of caffeine (which he had synthesized two years earlier), and Adolf Johann Friedrich Wilhelm von Baeyer (\$1905, died 1917, 0.8 ϵ) offered synthetic indigo on the market. The consequences

were not the same. The synthesis of caffeine had no effect on coffee plantations, whereas the Badische Anilin- und Sodafabrik, having found out how to manufacture artificial indigo in quantity by exploiting Baeyer's process, drove the world's indigo planters out of business.

This year marks a centennial for small killers. A hundred years ago, Emile van Ermengem found the bacillus of botulism. Masanori Ogata implicated rat fleas in

the cycle of bubonic plague epidemics. Ronald Ross (born 1857, 1.4 ϵ , \$1902) enjoyed his "Mosquito Day" (20 August), when he observed in the stomach walls of *Anopheles* mosquitoes black granules that pointed towards their role in the transmission of malaria. Kiyoshi Shiga discovered the dysentery bacillus. And the Dumdum Company of Calcutta introduced its not-so-magic bullet. On a larger scale, contemporaries could enjoy Emile Durkheim's *Suicide*, an exemplary application of statistics to sociology, and witness the almost simultaneous appearance of Charles Parsons' turbine-powered *Turbinia* among English warships and Ludwig Obry's precision torpedo.

None of the following died of any of the preceding, yet die they did in 1897: Edward Drinker Cope, American palaeontologist and proponent of the inheritance of acquired characteristics; Galileo Ferraris, Italian electrical engineer (born 1847, 1.5 ϵ);

Gut reaction: stomach wall of an *Anopheles* mosquito infected with malarial parasites.

Science Museum/Science and Society Picture Library

London School of Hygiene/SPL

Alfred Marshall Mayer, whose famous demonstration of the effect of a magnetic field on magnetic needles swimming on water provided an illustration of Thomson's theory of the arrangement of corpuscles in atoms; Victor Meyer, the world's greatest expert on nitro compounds; Fritz Johann Friedrich Theodor Müller, the German Darwinian who created the concept of Müllerian mimicry; Julius von Sachs, German botanist and historian of botany; Leonhard Sohncke, German crystallographer; and James Joseph Sylvester, English-Jewish mathematician notable for his work on invariants and for pinking a few anti-semitic undergraduates with his sword cane.

On the other side of the ledger, at least three people who were to win Nobel prizes for work in physics during the 1930s were born in 1897: Patrick Blackett, co-discoverer of the positron; John Douglas Cockcroft, co-inventor of the first particle accelerator to split atoms; and Irène Joliot-Curie, co-discoverer of artificially induced radioactivity. Anniversarilists in search of a bargain could do no better than celebrate these three centennials in one.

1847: Common currency

The philosopher has tried since antiquity to spy the constant and enduring under the flux that nature presents to the eye. The mid-nineteenth century saw the discovery of examples of transformations between various forms of power or force (to use the vocabulary of the time), such as electricity and magnetism, heat and mechanical action, and chemical processes and physical work. Several people guessed during the 1840s that these exchanges take place at invariable rates and can be reduced to a common currency. The most celebrated of these guesses was by Hermann von Helmholtz, who identified the common currency, the euro of physics and chemistry, as mechanical energy. He published the theory in 1847 in *Über die Erhaltung der Kraft*.

Another leading contributor to the estab-

lishment of fixed exchanges in 1847 was James Prescott Joule. Working in his father's brewery with the help of burly porters and refined thermometers, he gave out measurements of the exchange rate ('Joule's constant') for the transformation of mechanical work into heat.

Joule's was not the only exemplary measurement of heat made that year. There was the curious finding of François Hippolyte Walferdin that the temperature of the sea decreases by 1 °C for every 32.3 metres depth; and the disclosure by John William Draper (who later made himself famous by a silly book about the conflict between science and religion) that bodies when heated first give off red rays when they reach 525 °C, and become white hot between 1,200 and 1,300 °C. More heat than light may have been generated by Richard March Hoe's invention of the rotary press, which allowed the launch of mass-circulation newspapers.

There were some useful inventions too: the modern flute, by Theobald Böhm, an instrument maker in Munich; the sewing machine, by Elias Howe (died 1867, 1¢); steel razor blades, by the Dittmar company of Heilbronn; silver amalgam fillings, by Thomas Wiltberger Evans (died 1897, 1¢), an American dentist who lived in France; and chloroform in surgery and obstetrics, by James Young Simpson of Edinburgh. The year 1847 also saw the first use of adhesive postage stamps, in the United States.

Anaesthesia eased the sufferings of patients and made possible procedures that could not have been considered without it. But, by extending the time and type of operations, it also improved opportunities for infection. In those days, most surgeons considered washing a waste of time. When, in 1847, Ignaz Philipp Semmelweis insisted that students and doctors at the *Allgemeines Krankenhaus* in Vienna wash their hands between patients, the incidence of puerperal fever fell dramatically.

Other astonishing discoveries that year included Henri Victor Regnault's revelation that not even the permanent gases satisfy perfectly the law of perfect gases, and George Boole's demonstration, in his *Mathematical Analysis of Logic*, that deviations from standard logic and algebra may be rigorous and useful.

There are three more discoveries deserving of sesquicentennial attention. Gustav Robert Kirchhoff (died 1887, 1.1¢) enumerated the laws of the distribution of currents in electric circuits that immortalized his name. Karl Friedrich Wilhelm Ludwig, one of the creators of modern physiology, introduced graphical methods into his science with the kymograph, which could continuously record blood pressure and a host of other variables. And John Herschel published his *Results of Observations at the Cape of Good Hope*, completing the survey of the southern skies begun by Edmond Halley.

Sesquicentennial deaths include those of Alexandre Brongniart, the French geologist

Stitcher in time: the Howe sewing machine.

famous for his determination of the order of strata around Paris, and Georg Wilhem Muncke, opponent of *Naturphilosophie* and one of the editors of the 11-volume *Physikalisches Wörterbuch*. But the deaths cannot compete in number or distinction with the sesquicentennial births. These include the great inventors Alexander Graham Bell and Thomas Alva Edison, who should be given a joint and jolly anniversary; Edwin Ray Lankester, English naturalist and spokesman for science; Joseph Achille Le Bel, contributor to the theory of chemical structure; the Russian mathematicians Nikolay Egorovich Zhukovsky, 'the father of Russian aviation', and Egor Ivanovich Zolotarev, 'the most gifted member of the St Petersburg school of mathematics'; and the United Kingdom's Institution of Mechanical Engineers, the American Medical Association and the American Association for the Advancement of Science (AAAS). Maria Mitchell, the first woman member of the AAAS, discovered her first and only comet in 1847.

1797: Auntie's comets

Feminists and cometologists can join forces this year in celebrating the bicentenary of Heinrich Wilhelm Matthius Olbers' method of calculating cometary trajectories, which, now computerized, may alert us to the inevitable collision of Earth with a hairy star at a time now unspecified; and the discovery, by Caroline Herschel (aunt of John Herschel; see 1847), of her eighth comet in 11 years.

People positive and negative toward mathematics can join in celebrating some of the more peculiar productions of the discipline: Joseph Louis Lagrange, in his profound *Théorie des Fonctions Analytiques*, claimed to have established the differential calculus without the concept of the infinitesimal; and Lorenzo Mascheroni, in his *Geometria del Compasso*, showed how to do all the theorems and constructions in Euclid with a compass alone, so rediscovering the perverse invention of the Danish mathematician Georg Mohr (died 1697, 3¢).

It perhaps belongs to the same order of

Work shift: Joule's apparatus to measure transformation of mechanical work into heat.

ideas that the now standard geometrical representation of complex numbers, invented by Caspar Wessell in this same year 1797, was rejected 50 years later (1.5¢) by a much greater mathematician, Augustin Louis Cauchy (died 1857, 1.4 ¢). Cauchy wanted to do away with complex numbers altogether.

We cannot withhold any longer the information that 200 years ago Josiah Spode II, of bone-china fame, took over the family works from his father; John Fitch applied the screw propeller to a steamboat; William Wollaston discovered uric acid in gouty joints; and, to add to the ills of the human race, Cuba gave birth to cigarettes.

Other noteworthy births in 1797 included those of Joseph Henry, the American natural philosopher who became the first head of the Smithsonian Institution of Washington; Adrien Henri Laurent de Jussieu, the last of that ilk, which ran the Musée d'Histoire Naturelle in Paris for several generations and developed a 'natural' system of botanical classification in opposition to the 'artificial' scheme of Linnaeus; Carl Gustav Mosander, the Swedish chemist who discovered lanthanum; and the French mathematical physicists Jean Marie Constant Duhamel, the author of Duhamel's principle in partial differential equations, and the richly named Adhémar Jean Claude Barré Saint-Venant.

The bicentennial of the birth of Charles Lyell deserves special notice, for it offers a particularly clear case of the transmigration of souls. Lyell became the spokesman for uniformitarian geology through receiving the soul of James Hutton, the inventor of uniformitarian geology, who died the year Lyell was born. The case can be compared with Galileo-Newton, as explained five years ago (see *Nature* 355, 11; 1992).

1747: The Earth nods

Anniversarialists will certainly give a nod to astronomy this year. A quarter of a millennium ago James Bradley, the Astronomer Royal of England, demonstrated that, as Newton had suspected, the Earth's axis nutates, that is, nods, obeying the gravitational attraction of the planets. And Leonhard Euler suggested how to eliminate chromatic aberration from refracting telescopes, a contrivance that Newton had deemed impossible.

Because optics was associated so closely with astronomy in the eighteenth century, it would not be amiss to include in this set of anniversaries the extravaganza of Georges Louis Leclerc, Comte de Buffon (born 1707, 2.9¢), who ignited a piece of wood at a distance of 47 metres with a mirror made of more than 200 panes of glass.

Two of the few instances of science-based technology in the eighteenth century had their origins in 1747. The German chemist Andreas Sigismund Marggraf, searching for a plant and process to give the northern countries their own source of sugar, showed the promise of beet root.

The second significant bit of sci-tech derived from the retirement pursuits of the American printer Benjamin Franklin. While amusing himself with experiments on electricity, he hit on the concepts of electricity plus and minus. That was in 1747. Within two years he had identified the power of points in discharging insulated electrified bodies at a distance and, in one of the bold analogies characteristic of his thought, compared the bodies with clouds and the discharges with lightning. He suggested that an elevated, grounded point, aimed at the heavens, could despoil storm clouds of their destructive power, and so invented the lightning rod. As for his

chemistry during the eighteenth century. It supposed that metals contained a special principle that gave them their peculiar lustre and ductility, and explained, by its loss, their calcination and oxidation. The theory was introduced in or about 1697 by Georg Ernst Stahl, professor at the University of Halle. To personalize the achievement of Semmelweis (see 1847), we record that Stahl's wife died of puerperal fever the year before phlogiston was born.

Mathematicians of the time made a practice of challenging one another to solve tough problems. A famous example, because of the calibre of the competitors and the fecundity of the solutions, was set by Johann Bernoulli I. He asked for the shape of the curve along which a body will descend between two given points under the force of gravity. Newton and Leibniz supplied solutions to this problem of the brachistochrone, which, being more elegant than Bernoulli's, pleased posterity but annoyed him.

It was not a good year for other mathematicians, either. Stefano degli Angeli, who had taken part in the debate over the Earth's motion, died, as did Georg Mohr (see 1797).

1647: Mountains of the Moon

Commemorators of the days on which even great people seldom accomplish anything, that is, the days they die, have two strong disciples of Galileo to celebrate. They are Bonaventura Cavalieri, developer of a forerunner of the calculus, who taught the Copernican system after its condemnation while a professor at the University of Bologna within the Papal States; and Evangelista Torricelli, also a mathematician, but better known for his invention of the mercury barometer and the demonstration of a space void of air. Torricelli's insights were developed with great virtuosity by Blaise Pascal, in his *Expériences Nouvelles Touchant le Vide*, also of 1647.

There is no doubt that Johann Hevelius's *Selenographia* deserves an anniversary celebration. The product of the substantial wealth and good eyesight of this Danzig brewer, the *Selenographia* established many features of lunar geography and introduced the nomenclature of mountains and maria still used by Moon men.

1597: The Pope's satellites

The pick of the year is Andreas Libavius's *Alchemia*, which emphasized the importance of chemistry for medicine and the importance of clear writing for chemistry. Perhaps the fact that his first university position was a professorship of history accounts for his insistence on clarity, comprehensibility and practicality. He was also a medical doctor, logician, theologian and pamphleteer.

Notable birthdays in 1597 were those of Henry Gellibrand, professor at Gresham College in London and a discoverer of the secular change in the Earth's magnetic field;

Moon gazer: from Hevelius's *Selenographia*.

single-fluid theory of electricity, which flowed from his concept of positive and negative charge (or vice versa), it was to find full confirmation in our discovery of the year (see 1897).

Three more innovations full of the future are strong candidates for 2.5 centennials this year. James Lind's demonstration that citrus fruits cure scurvy argued for the introduction of lime juice among the rations of the British sailor, thereby and thereafter known as a 'limey'. This life-saving measure was discontinued when progress brought tinned juices — which unfortunately were rendered ineffectual by loss of their vitamin C during canning. The business was not cleared up before the end of the nineteenth century. The launching of the *Encyclopédie* of Denis Diderot and Jean Le Rond d'Alembert, a vast compendium of knowledge and grand instrument of enlightenment, has been credited with subverting the *ancien régime* and setting the course of modernity. The establishment of the École des Ponts et Chaussées, the French state school of civil engineering, may be regarded as the foundation stone of all modern technical higher schools, whose graduates have transformed the world.

1697: Elegant curve

The often maligned theory of phlogiston brought order into an important branch of

Raffaello Magiotti, a student of Galileo's who also worked with Torricelli (see 1647); Giambattista Odierna, another Galileian, an observer of comets and insects; and Anton Maria Schyrllaes de Rheita, a Bohemian astronomer, one of the early mappers of the Moon and the hopeful dedicator of spurious satellites of Saturn to Pope Urban VIII.

1547: Healing power of grass

We believe that in 1547 the Dominican Georges Nernard Penot published his *De Aquae Naturalis Virtute*, based on the medical merits of walking with bare feet on wet grass; and that death took the accomplished astronomer Johannes Schöner (born 1477, 5.2¢), a Catholic priest turned Lutheran, who printed his astronomical works on his own press, made his own globes and edited the manuscripts of Regiomontanus.

1497: No smoke without harm

The first description of the tobacco plant was given 500 years ago by Romano Pané, a monk who sailed on Christopher Columbus's second voyage to America. Pané called the plant *Herba inebrians*. It did not take a surgeon general to discover its effects on humans.

While on the subject of travel, in 1497 John Cabot reached Newfoundland and Vasco da Gama rounded the Cape of Good Hope on the way to India.

1397: Florentine fete

Florence should consider staging a first-class intellectual orgy in honour of the six-hundredth anniversary of the birth of the extraordinary Doctor Paolo del Posco Toscanelli, friend and adviser of the architects Leone Battista Alberti and Filippo Brunelleschi, and encouraging correspondent of Columbus.

1247: Bethlehem birth

Recent events associated with St Bartholomew's Hospital in London remind us that not even venerable institutions are immortal. But we are not privy to any plans to close the psychiatric hospital, once in London but now in Beckenham, Kent, that contributed a shortened form of its name, Bedlam, to popular culture. Its parent, the Priory of St Mary of Bethlehem, was founded seven-and-a-half centuries ago.

1947: Rays of hope

We do not know of anyone active in 1947 who was associated with so many world-historical events as Benjamin Franklin or Paolo Toscanelli. Still, in the aggregate, people managed to accomplish more than a little that year.

The physicists at the University of California at Berkeley got their 32 MeV proton linear accelerator going under the leadership of Luis Alvarez (\$1968). Their colleagues at General Electric first detected synchrotron radiation from their electron accelerator. Their colleagues at Bell Labs,

John Bardeen (\$1956) and Walter Houser Brattain (\$1956, died 1987, 0.1¢), found the transistor effect. In England, Cecil Frank Powell (\$1950) caught a cosmic ray in the act of changing from one meson to another, so confirming the theory that there exist at least two mesons, one an intermediate in nuclear forces, the other rather useless. Again in England, Dennis Gabor (\$1971) developed the basic concept of holography. In the United States, Willis Eugene Lamb (\$1955) announced the 'Lamb shift', a key phenomenon in quantum electrodynamics.

Piped pollution: it was a monk who gave Europe its first news of the tobacco plant.

And, in Australia, John C. Bolton made the first discovery of an optical radio source, in the Crab Nebula.

The Dead Sea Scrolls were discovered and dated with the help of the then novel analysis of their content of radioactive carbon. The method later earned its inventor, Willard Frank Libby, a Nobel prize (\$1960).

Old Toscanelli, and Columbus too, would have been interested to see a C-54 transport plane soar across the Atlantic without a crew, entirely on automatic pilot; and they would have appreciated that the flight of the first aeroplane to move at supersonic speed presaged that those who could afford it could cross the ocean about 150 times faster than Columbus did. To provide meals at a corresponding tempo, Raytheon brought radar from the skies to the kitchen in the first microwave oven, which was not a commercial success.

1947 ± 25: Piped muzak

We begin with two of the worst ideas of the twentieth century. In 1922, George Squire patented a method for sending music over wires, which led to the Muzak Corporation, and thence to the piped music now the scourge of everyday life. Fifty years later, in an unrelated act, the Board of Education of the State of California decreed that the Biblical account of creation should be given equal time with Darwinian evolution in schools. This decision should be celebrated

with Alexander Ivanovich Oparin's studies of the naturalistic origins of life, which began in 1922; and Muzak should be celebrated along with the compact disc, invented in 1972.

We have more vitamins on offer (see 1747). In 1922, Herbert Evans and Katherine Bishop announced "a hitherto unrecognized dietary factor essential for reproduction" (vitamin E), while Elmer McCollum isolated the anti-rickets substance vitamin D. That these three scientists achieved an average age of 88 might justify a modest advertising campaign. Still on the subject of metabolism, Gerty and Carl Cori settled in the United States in the same year, and began their work on glycogen and glucose that a quarter of a century later made them the third husband-and-wife team to share a Nobel prize (\$1947).

In 1922, Leonard Woolley began to excavate Ur and a team led by the Earl of Carnarvon and Howard Carter opened Tutankhamen's burial chamber. Fifty years later, Murray Gell-Mann (\$1969) entered further into the mysteries of life with his anthropomorphic coloured and tasty quarks, the basis of quantum chromodynamics. Bodies also became more visible in 1972, when Godfrey Hounsfield (\$1979) presented the results of the first computerized axial tomography scan of the human brain, produced in a Wimbledon hospital the year before.

In 1922, an aeroplane landed for the first time on an aircraft carrier, and Aleksander Aleksandrovich Friedman showed that the theory of relativity allowed for an expanding Universe. In 1972, men landed for the fifth and sixth times on the Moon and the US spaceship Pioneer 10 left the Solar System, the first man-made object to do so, in the never-ending quest and test of the limits of knowledge. Still, in our opinion, the first choice for a quarter-century anniversary is the passing of the Clean Air Act by the US Congress, which, among many other consequences, has helped in the drive to remove the smoke of *Herba inebrians* (see 1497 and 1797) from public places.

Where we came in

Finally, first among the least significant anniversaries recommended for commemoration this year is the one-tenth centennial of the first appearance of this column and the half-centennial of the pioneering work of Edgar C. Smith in "Scientific centenaries in 1947" (*Nature* 159, 16-18; 1947). Writing at a primitive stage of the anniversarial art, Smith offered only relevant facts and sound information; as faithful readers know, we have been able to transcend these limitations. □

J. L. Heilbron is at Ofslang House, Mill Street, Islip, Oxfordshire OX5 2SZ, UK. W. F. Bynum is at the Wellcome Institute for the History of Medicine, 183 Euston Road, London NW1 2BE, UK.

Quantum security is spookily certain

Richard J. Hughes

Quantum cryptography has been suggested as a way for two people to communicate in perfect secrecy. Now it has been shown practically that more than one listener can be brought in.

THE representation of information by classical physical quantities, such as the voltage levels in a microprocessor, is familiar to everyone. But over the past decade the science of quantum information has been developed, which describes binary information in the form of two-state quantum systems such as photon polarization states. Remarkable new capabilities have been predicted that make use of quantum-mechanical superpositions of information, and one of these, quantum cryptography¹, is now close to becoming practical. On page 47 of this issue², Paul Townsend reports the first realization of a multi-user version of quantum cryptography.

One of the main goals of cryptography is for two parties ('Alice' and 'Bob') to render their communications unintelligible to a third party ('Eve'). This can be accomplished if Alice and Bob both possess a secret random bit sequence, known as a cryptographic key. For example, in 'one-time pad' encryption Alice adds the key (modulo two, if they are using binary data) to the original message, known as plaintext, and communicates the sum (ciphertext) to Bob. He can recover the plaintext by adding his key (modulo two) to the ciphertext, but Eve, who has monitored the transmitted ciphertext, is unable to discern the underlying plaintext through the randomization introduced by Alice's key. So, although key material conveys no useful information itself, it is a very valuable commodity, and methods for Alice and Bob to generate key material securely are correspondingly important.

Using quantum cryptography, or more accurately, quantum key distribution (QKD), Alice and Bob can create shared cryptographic key material whose security is assured by the laws of quantum mechanics³⁻⁶. They independently generate secret random number sequences, which then undergo a bit-wise comparison: Alice prepares and transmits a *single* photon for each bit, and Bob measures it. Alice's photon state preparations and Bob's measurements are determined by their bit values, and are chosen from sets of non-

orthogonal possibilities, such as linear and circular polarization. This comparison algorithm, which may be publicly known, ensures that Bob detects a photon (with some quantum-mechanically determined probability) only if he has the same bit value as Alice. They retain only the de-

germanium or indium-gallium-arsenide avalanche photodiodes can be persuaded to detect single photons, but at the penalty of high noise and hence a high error rate. Removing these errors reduces the rate of key transmission and limits distances to 100 km or so. Optical amplifiers cannot be used to extend this range because they cannot replicate the non-orthogonal quantum states used in QKD.

Townsend's new results show that optical-fibre QKD could be operated not only in a one-to-one configuration but also in a passive one-to-many mode. Specifically, each photon that Alice sends out makes a random choice at a fibre splitter of travelling to precisely one of several receivers (different Bobs). Then each Bob responds to Alice informing her of which photon in her sending sequence he received. The indivisibility of the photon allows Alice to generate an independent key with each Bob. Thus, QKD could be used to generate cryptographic keys over 'open' optical fibre links between secure 'islands' separated by distances such as London-Cheltenham in the United Kingdom, or in the United States between different government agencies in the Washington DC area.

In other experiments¹³, QKD is being developed for free-space, line-of-sight communications, including possible re-keying of satellites in low-Earth orbits. At present, though, transmission distances through the atmosphere are only tens of metres.

Cryptographic requirements also limit potential applications. Because QKD has the ultimate security assurance of a law of nature, its inventors proposed that it should be teamed up with one-time pad

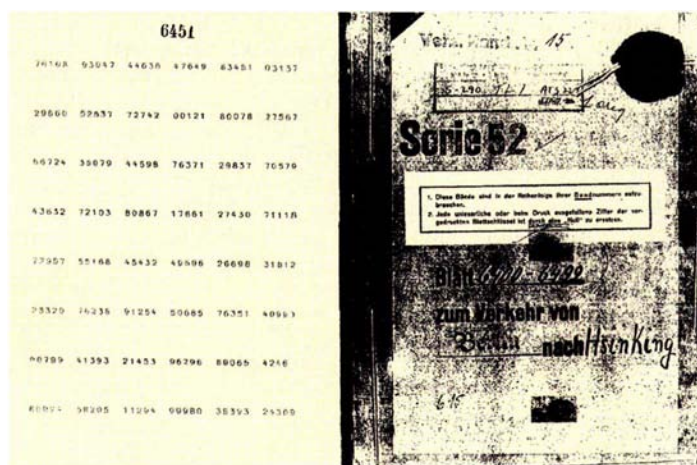


FIG. 1 A 'one-time pad' of random numbers used for encrypting communications between Germany and China during the Second World War. (Courtesy of the National Cryptological Museum, Fort Meade, Maryland, USA.)

tested bits from their initial sequences. These subsets are the raw key material from which a pure key is distilled using classical error-detection techniques, to eliminate the effects of noise.

Eve can neither 'tap' the key transmissions, owing to the indivisibility of a photon, nor copy them, owing to the quantum 'no-cloning' theorem. Furthermore, the non-orthogonal nature of the quantum states ensures that if Eve makes her own measurements she will be detected through the increased error rate arising from the irreversible collapse of the wavefunction that she introduces.

QKD is possible over optical-fibre paths many kilometres long: the necessary quantum coherence of the QKD transmissions persists even outside the controlled environment of a physics laboratory⁷⁻¹². At the infrared wavelengths required,



FIG. 2 A very simple cipher from the Sherlock Holmes story "The adventure of the dancing men": a stick figure (a dancing man) is substituted for each letter of a message.

encryption³⁻⁵ to provide provably secure cryptography on demand. Although theoretically appealing, this union is impractical for many applications because one-time pad encryption requires a key as long as the message — and, with current technology, key generation rates are only in the kHz to MHz range.

But Alice and Bob could use short quantum keys in other (symmetric key) systems. For example, they might use a few thousand secret bits as the 'seed' for synchronized, cryptographically secure pseudo-random number generators. Used in this way, QKD would still offer many advantages over existing key distribution methods in terms of security and ease of use.

Traditional key distribution, using trusted couriers, requires cumbersome security procedures for preparing, transporting and handling the key before any communications can take place, and in some cases is impractical — one can't send a courier to re-key a satellite. In contrast, quantum keys do not even exist before the QKD transmissions are made, and a key can be generated when the message is transmitted.

Public-key cryptography also avoids many of the difficulties of key distribution by courier, but provides only the conditional security of 'intractable' mathematical problems, such as integer factorization. For example, Bob publishes a public key (an integer exponent, e , and an integer modulus, N) and keeps secret a private key, which is essentially the prime factorization of N . Alice encrypts her transmission to Bob by taking the e th power of her data and sending him the remainder modulo N . Standard number theory results allow Bob to decrypt Alice's communication because of his knowledge of the prime factors of N , but Eve would have to factorize N herself, and N is chosen to make this computationally intractable. But it is notoriously difficult to assess accurately an adversary's computing power over the useful lifetime of encrypted information, which may be years or even decades. Unanticipated advances in fields such as quantum compu-

tation could render public-key methods not just insecure in the future but also retroactively vulnerable. So QKD could be used for real-time key generation in cryptographic applications where this long-term risk is unacceptable.

Quantum cryptography is likely to be the first practical application of the foundations of quantum mechanics, and it illustrates the often unexpected value of basic research. Like any other cryptographic system, QKD will be subject to stringent testing and evaluation before

being accepted, and physicists should not be surprised if the cryptographic community questions the fundamental principles that are at its heart. But perhaps the rigorous scrutiny required for cryptographic evaluation will inspire new insights into the strange world of quantum physics. □

Richard J. Hughes is in the Physics Division, Los Alamos National Laboratory, University of California, P-23, MS H803, Los Alamos, New Mexico 87545, USA.

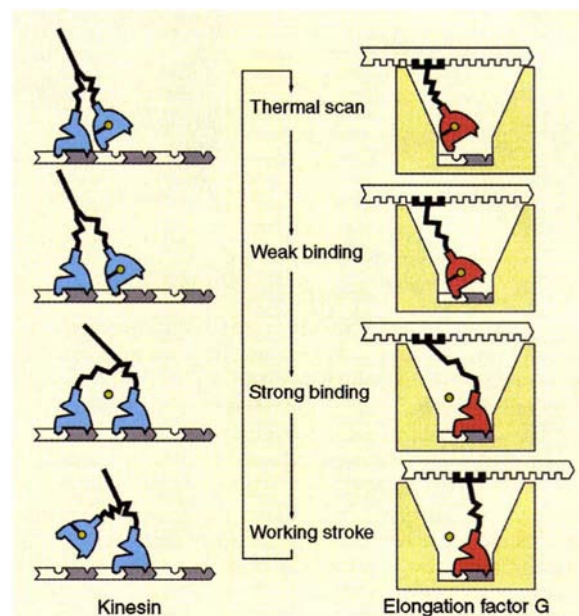
MECHANOCHEMISTRY

A protein-making motor protein

R. A. Cross

RIBOSOMES are programmable molecular devices that synthesize proteins. They work like laboratory peptide synthesizers, by robotic repetition of a synthetic cycle. But the molecular-scale ribosomal machinery is considerably — about 10^5 -fold — faster, and more accurate, than the laboratory model. The ribosome binds its

tions of this cycle progressively extrude the growing protein chain. On page 37 of this issue, Rodnina *et al.*¹ argue that the ribosomal protein that catalyses the mRNA translocation step — elongation factor G (EF-G) — does not simply bind and dissociate once per notch, but that it is an authentic molecular motor.



Apparent mechanochemical parallels between elongation factor G (EF-G) and cytoskeletal motors. Both kinesin (left) and EF-G (right) consist of a nucleotidase domain coupled to a sprung actuator arm. In this moderately speculative scheme, the ratcheting of kinesin's cargo is coupled to ADP (yellow circle) release. Rodnina *et al.*¹ show that GTP hydrolysis by EF-G is too fast to be coupled directly to mRNA ratcheting on the ribosome tape, suggesting that this process also is coupled to product release.

punched-tape set of instructions, the linear messenger RNA code, and sequentially reads each instruction. It then selects the appropriate monomer (EF-Tu-aminoacyl-transfer RNA complex) from the 20 possibilities in the surrounding solution, plugs it on, and translocates one notch (or codon) along the instruction tape. Itera-

What must an enzyme do to qualify as a molecular motor? All enzymes are mechanical, as well as chemical, devices. Events in the active site tend to induce (albeit usually modest) shifts in the global shape of the molecule, exactly because the active site is mechanically coupled to the rest of the protein structure. Motors are enzymes that can amplify and harness such cyclic shape changes² so as to step themselves along 'tracks' such as actin, microtubules and DNA. All of the motors that have been described so far use the turnover of nucleotides such as ATP or GTP to effect conformational changes and to drive correlated binding to and release from the track. So to understand how a particular molecular motor works, we need to analyse the mechanisms by which its active-site chemistry and its global mechanics are coupled together³.

Rodnina *et al.*¹ have explored the coupling between nucleotide turnover and the EF-G-catalysed ratcheting of the mRNA tape relative to the ribosome. The experiments set out to test the established notion that GTP hydrolysis is directly coupled to the release of EF-G from the mRNA on the ribosome. Translocation of the mRNA tape is the

1. Hughes, R. J. *et al. Contemp. Phys.* **36**, 149–163 (1995).
2. Townsend, P. D. *Nature* **385**, 47–49 (1997).
3. Bennett, C. H. & Brassard, G. *Proc. IEEE Int. Conf. Computers, Systems and Signal Processing, Bangalore* 175–179 (IEEE, New York, 1984).
4. Bennett, C. H. *Phys. Rev. Lett.* **68**, 3121–3124 (1992).
5. Bennett, C. H. *et al. J. Cryptogr.* **5**, 3–23 (1992).
6. Ekert, A. K. *Phys. Rev. Lett.* **67**, 661–663 (1991).
7. Townsend, P. D. *Electron. Lett.* **30**, 809–811 (1994).
8. Marand, C. & Townsend, P. D. *Opt. Lett.* **20**, 1695–1697 (1995).
9. Franson, J. D. & Ilves, H. *Appl. Opt.* **33**, 2949 (1994).
10. Muller, A., Zbinden, H. & Gisin, N. *Nature* **378**, 449 (1995).
11. Muller, A. *et al. Europhys. Lett.* **33**, 335 (1996).
12. Hughes, R. J. *et al. Lecture Notes in Computer Science* Vol. 1109, 329–342 (Springer, New York, 1996).
13. Jacobs, B. C. & Franson, J. D. *Opt. Lett.* **21**, 1854–1856 (1996).

last step in the ribosome catalytic cycle, and each such cycle of addition of one amino-acid residue to the growing protein takes about 50 milliseconds. The ribosome spends most of this time sampling aminoacyl-tRNA adducts from the surrounding solution, searching for a fit between the current instruction (the codon) and an incoming amino acid. Incoming aminoacyl-tRNAs are complexed to another GTPase elongation factor called EF-Tu. Once a base-paired fit is established, EF-Tu shuttles the amino acid to an active centre that is occupied by the end of the growing protein, and then dissociates. EF-G then binds and does its job of ratcheting the tape.

Up until now, thinking about the mechanisms by which EF-G and EF-Tu might act has been coloured by the strong structural resemblance between the GTP-binding domains of the two elongation factors and the GTP-binding domain of the oncoprotein Ras. Ras is a switch or, more properly, a relay, that flicks between a GTP-bound 'on' state, in which it binds to its target, and a GDP-bound 'off' state, in which it dissociates, according to a hormonal input signal. So the aspects of EF-Tu- and EF-G-mediated catalysis that have been emphasized previously are the GTP-dependent binding and GDP-dependent release steps. But during the past year new crystal structures of EF-Tu and EF-G have been solved⁴⁻⁸, and they have revealed some big surprises.

First, the distal domains 4 and 5 of EF-G are seen to have the same shape and size as the tRNA binding partner of EF-Tu. This remarkable molecular mimicry suggests that the two elongation factors work alternate shifts at the same binding site. Second, the transition from GTP- to GDP-binding in the EF-Tu active site causes sizeable movement of the attached tRNA, encouraging models⁹ in which the tRNA of EF-Tu and the mimetic domain 4 of EF-G act as mechanical manipulators — something like robotic arms.

So how is the putative mechanical action of EF-G coupled to the turnover of GTP? Rodnina *et al.*¹ have used single-turnover experiments (in which the event of interest occurs only once) to show that GTP hydrolysis occurs five times faster than mRNA ratcheting. In other words, hydrolysis precedes translocation by quite a considerable time. The consequences of

this finding are profound. It implies that the bond energy that is derived from the hydrolysis of GTP is stored in the protein, to be used later for mechanical work. The protein is like a compressed spring that is held against a trigger — ribosome-binding releases the trigger and the protein flies open, exerting force and allowing the phosphate that is generated as a result of GTP hydrolysis to exit from the active site.

There are striking parallels between this mode of chemical-mechanical coupling and the coupling that is used by the classical motor proteins kinesin, myosin and dynein (see figure). In these cytoskeletal motors, the product phosphate and/or the nucleoside diphosphate (ADP) are essentially trapped in the active site until the motor protein finds its binding site. The motor then adopts a strongly bound conformation, prior to its mechanical power stroke. Rodnina *et al.* show that mRNA translocation is inhibited by mutagenic removal of domain 4 — the part of EF-G that mimics the tRNA of EF-Tu, whereas ribosome-activated GTP hydrolysis is unaffected. The implication is that the motor has lost its robotic actuator arm. Domain 4

is also required for the dissociation and recycling of EF-G: deletion of domain 4 leaves the motor stranded in the ribosome site, able to perform a single round of GTP hydrolysis but unable to dissociate. So the mechanical motion of domain 4 seems to be necessary to drive the motor into its dissociating conformation.

EF-G also gets 'stuck' if non-hydrolysable GTP analogues are used, emphasizing the fact that mechanochemical coupling works both ways — just as active-site events can generate useful force, so active-site chemistry is sensitive to mechanical strain acting on the enzyme molecule. The challenge now, in all of these motor systems, is to set up experiments that measure two-dimensional rate constants for the linked mechanical and chemical equilibria, preferably at the single-molecule (single ribosome) level. In other words, we need to ask the question, "what happens to the active-site chemistry when I grab hold of this molecule and pull?" □

R. A. Cross is in the Molecular Motors Group, Marie Curie Research Institute, The Chart, Oxted, Surrey RH8 0TL, UK.

EARTHQUAKES

Long odds on prediction

Ian Main

EARTHQUAKES and horse-racing are both complex phenomena where the prediction of individual events is inherently difficult, if not impossible. No one expects the bookmaker's favourite to win every time, but everyone expects the bookie to live well off an astute calculation of the odds. But the perception of the public, the media and many government funding agencies is that we should nevertheless go on betting heavily on the prediction of individual earthquakes, rather than spending money on a careful calculation of the odds — that is, estimation of the seismic hazard.

A useful earthquake prediction requires specification in advance of the location, magnitude and time of an individual event, within narrow limits, otherwise a programmed evacuation could not take place^{1,2}. Prediction in this sense has proved elusive. In contrast, statistical estimates of the seismic hazard are based on a calculation of the likelihood of ground shaking from an understanding of the source mechanics of a population of earthquakes.

The thorny question of how to respond to the threat of earthquakes was at the core of a meeting* on the validation of schemes for earthquake prediction. Such schemes are sometimes based on a deterministic physical hypothesis, but more commonly on an empirical observation of

geophysical or geochemical precursory 'anomalies'. The most cited physical hypothesis was based on the observation of dilatancy (an increase in sample volume due to microcracking) and associated precursors observed in laboratory tests, proposed in the 1970s (ref. 3). But the predicted anomalies failed to materialize, and the hypothesis was rejected — that is, the results failed to scale linearly over the spatial and temporal scales that separate laboratory tests and large earthquakes.

At the meeting, consensus emerged on several points. First, given the dynamic complexity of earthquake sources and the material heterogeneity of the Earth, there are no clear reasons why the reliable prediction of individual earthquakes is possible⁴. Second, we are not yet in a position to identify any significant and unambiguous earthquake precursors, even with the benefit of hindsight⁵ ('past-posting' in betting terminology). Third, individual predictions, and prediction methods, should be stated so that their success or failure can be objectively and unambiguously determined. ▶

*Assessment of Schemes for Earthquake Prediction, Royal Astronomical Society/Joint Association for Geophysics Discussion Meeting, London, 7-8 November 1996. Abstracts can be viewed at http://www.seismo.demon.co.uk/Nov7th/second_circular.html or are available by e-mail from russ.evans@bgs.ac.uk.

1. Rodnina, M. V., Savelsbergh, A., Katunin, V. I. & Wintermeyer, W. *Nature* **385**, 37-41 (1997).
2. Block, S. M., *Cell* **87**, 151-157 (1996).
3. Vale, R. D. *J. Cell Biol.* **135**, 291-302 (1996).
4. Nissen, P. *et al. Science* **270**, 1464-1472 (1995).
5. Kawashima, T., Berthet-Colominas, C., Wulff, M., Cusack, S. & Leberman, R. *Nature* **379**, 511-518 (1996).
6. Polekhina, G. *et al. Structure* **4**, 1141-1151 (1996).
7. Abel, K., Yoder, M. D., Hilgenfeld, R. & Jurnak, F. *Structure* **4**, 1153-1159 (1996).
8. Al-Karadaghi, S., Evarsson, A., Garber, M., Zheltosova, J. & Liljas, A. *Structure* **4**, 555-565 (1996).
9. Abel, K. & Jurnak, F. *Structure* **4**, 229-238 (1996).

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Aftermath of the magnitude 8.1 Michoacan earthquake which devastated Mexico City, 400 km from the epicentre, in September 1985. On current knowledge, individual earthquakes cannot be predicted in any reliable way. Hazard mitigation, based on a lower-profile statistical approach, remains the better prospect.

Finally, in the absence of reliable prediction methods, we should concentrate on hazard mitigation based on a better understanding of earthquake source mechanisms, their statistical properties, the propagation of seismic waves and the response of individual sites, buildings and infrastructure to seismic vibration⁵. This lower-profile, statistical approach does not aim to give, say, a few hours' or days' warning for inhabitants of an area to leave it; rather, it results in guidelines for the construction of buildings and other structures that will withstand the forces earthquakes impose upon them. Most deaths and injuries from earthquakes are caused by collapse of buildings, or secondary effects such as ensuing fires, so this approach should save more lives, given the current state of knowledge^{6,7}.

Why have earthquakes proven so difficult or impossible to predict? There are two possibilities: either detectable and reliable empirical precursors do generally exist, but our instrumentation cannot measure them; or the physics of earthquakes is too sensitive to small fluctuations to produce reliable precursors. In fact, modern theories of earthquakes hold that they are critical, or self-organized critical, phenomena⁸, implying a system maintained permanently 'on the edge of chaos', with an inherent random element and 'avalanche' dynamics with a strong sensitivity to small stress perturbations. The notion of self-organized criticality is consistent both with the observed frequency-magnitude relation of earthquake

populations, and the presence of $1/f$ (power-law) noise in almost all borehole spectra on a variety of rock types in different tectonic areas (P. Leary, Univ. Edinburgh).

Evidence of 'fracture criticality' has been inferred from observations of seismic anisotropy, using a new theory of stress-induced, directionally dependent, contemporaneous crack opening and closure, which need not produce large-scale dilatancy (S. Crampin and S. Zatsepin, Univ. Edinburgh). Time-varying anisotropy due to stress changes may therefore be observable in the form of temporal changes in shear-wave polarization. This approach constitutes a middle course, where a full earthquake prediction is not possible, but the probability of occurrence of possible events is temporarily elevated above the long-term seismic hazard⁹.

According to one view, it is "highly unlikely" that reliable precursors exist (R. Geller, Univ. Tokyo). Indeed, no single precursor satisfying the validation criteria developed by the International Association for Seismology and Physics of the Earth's Interior has ever been observed unambiguously (D. Booth, British Geological Survey). These criteria include a precise definition of the anomaly, an explicit statement of the signal-to-noise ratio, detection at more than one site, and a full publication of negative as well as positive evidence⁵.

There are many reported 'anomalous' electrical signals in earthquake zones, recorded as voltage oscillations between two electrodes coupled to the ground and separated by relatively long distances. Case studies of such observations were presented for Japan (Y. Enomoto, Mechanical Engineering Laboratory, Tsukuba, Japan) and Crete (F. Vallianatos, Chania, Crete). However, as in other descriptions of the same techniques, there is no clear-cut statistically significant correlation with individual earthquakes during the recording period.

In analysing such signals, it is also essential to be able to identify a plausible physical meaning for them — for instance, assignment of many electric precursors to events preceding an earthquake, notably by P. Varotsos *et al.* (the VAN group) in Greece, violates the principle of energy conservation (P. Bernard, Inst. Physique du Globe, Paris). It is an indication of mistaken priorities that earthquake 'prediction' as advocated by this group absorbs more funding than research programmes to improve building design practice in Greece (S. Stiros, Inst. Geology and Mineral Exploration, Athens).

Vague 'predictions', with loose limits on

their magnitude, and time and place of occurrence, can also give spurious success rates, even for a purely random process (Bernard). This theme was taken up by several advocates of rigorous statistical testing and evaluation of the significance of reported precursors and predictions (Y. Kagan, Univ. California, Los Angeles; F. Mulargia, Univ. Bologna; P. Stark, Univ. California, Berkeley). For example, a prediction scheme should perform better than a naive rule, such as predicting that large events will have aftershocks (Stark, Mulargia). When this is done, with full access to the facts (successes — hits; and failures — misses or false alarms), it is hard to be optimistic about the prospects for reliable prediction. In contrast, methods already exist for assessing long-term hazard from combining instrumental, historical, geological and geodetic data (Y. Kagan, ref. 10).

Why is it hard to get this 'negative' conclusion, based on our experience of prediction, across? Speaking with the conviction of a prophet outside his adopted country of Japan, Geller laid into the sloppy practice and publicity-seeking activities of a minority of scientists and even non-scientists. We have had more than 100 years of failure in attempts to predict individual earthquakes, and "the public, media and government authorities must be clearly informed that earthquake prediction in its popularly understood sense is impossible at present, that all attempts to predict earthquakes to date have been failures, and that there are no reasonable prospects for prediction in the near future".

Given the current state of knowledge, the best bet with a guarantee of return in reducing the threat from earthquakes is on hazard mitigation. Even then, it is likely that there will continue to be some failures. The next best bet, perhaps in the form of an each-way flutter, seems to be on the establishment of the possibility of a seismic hazard that may be time-dependent. In the absence of past-posting, the prediction of individual earthquakes remains a long shot. □

Ian Main is in the Department of Geology and Geophysics, Grant Institute, University of Edinburgh, West Mains Road, Edinburgh EH9 3JW, UK.

1. Wood, H. O. & Gutenberg, B. *Science* **82**, 219–220 (1935).
2. Macelwane, J. B. *Bull. Seismol. Soc. Am.* **36**, 1–4 (1946).
3. Scholz, C. H., Sykes, L. R. & Aggarwal, Y. P. *Science* **181**, 803–810 (1973).
4. Main, I. G. *Rev. Geophys.* **34**, 433–462 (1996).
5. Wyss, M. (ed.) *Evaluation of Proposed Earthquake Precursors* (Am. Geophys. Un., Washington DC, 1991).
6. Reiter, L. *Earthquake Hazard Analysis* (Columbia Univ. Press, New York, 1991).
7. Press, F. *et al.* *Confronting Natural Disasters* (Nat'l Acad. Press, Washington DC, 1987).
8. Bak, P. & Tang, C. J. *Geophys. Res.* **94**, 15635–15637 (1989).
9. Keilis-Borok, V. I. & Gossobov, V. G. *Phys. Earth Planet. Inter.* **61**, 73–83 (1990).
10. Main, I. G. *Nature* **357**, 27–28 (1992).

Squirt smugly, scallop!

Steven Vogel

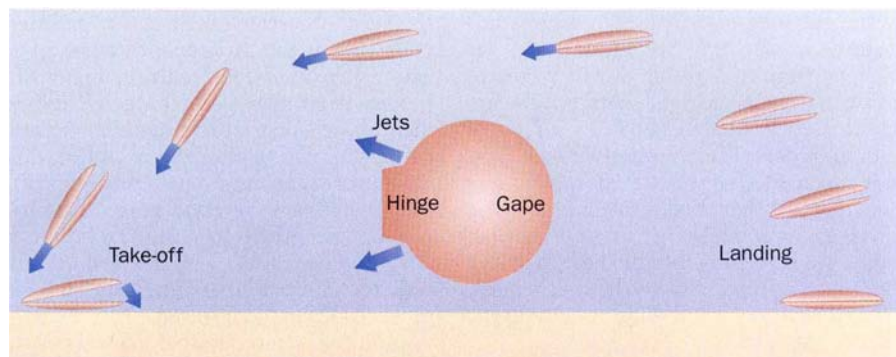
BIVALVE molluscs are not famous swimmers. But several kinds, most notably scallops, can swim and do so by repeatedly clapping the halves of their shells together. Some approach a speed of three kilometres per hour, although briefly — a bout of swimming lasts a few seconds and covers a few metres at most — and only when threatened by a predator, such as a starfish. Counter-intuitively, a scallop moves gape forward and hinge aft, as shown in the figure. At the same time as it obstructs the gape with a curtain of soft tissue, the scallop closes its shell and squirts water through a pair of apertures that flank the hinge. The shell is closed, and held closed, by a large, central adductor muscle (which is normally found denatured and covered in sauce on a dinner plate). When this muscle relaxes, a compressed elastic pad near the hinge reopens the shell.

Various aspects of scallop swimming have been studied over the past few decades: the lift of the shell, the resilience of the hinge pad, the performance of the muscle and how this propulsive system changes during growth. Now two papers from Edwin DeMont's group, published in the *Canadian Journal of Zoology*¹ and the *Journal of Experimental Biology*², allow us to form a good picture of the whole system.

Such a complete view of the mechanics of a locomotor system is rare, and more surprising still is the suitability of the scallop as the subject of such an analysis. In most organisms, overall motions are more complex, external forces vary less regularly over time, mechanical components are not as linearly elastic, muscle shortening is less regular, or power losses are harder to measure.

The adductor muscle and the elastic hinge element of a scallop run directly and independently from one half-shell to the other. The two half-shells move in a smooth cycle that directly mirrors the strains on both the muscle and the hinge pad. By measuring these movements and the overall motion through the water, the speed and path of swimming and the rate at which water is expelled can be determined. Consistent with this, Cheng *et al.*^{1,2} have shown that neither the way in which movements of the shell power the jet nor the forces resulting from motion through the water present severe analytical difficulties. But is the scallop as crude a swimmer as its clumsy-looking performance suggests? Or is it a finely tuned, if unlikely looking, design? By considering the precise mechanisms by which scallops swim, the authors argue strongly in favour of the latter case.

The least energy-expensive way to move to and fro is to store energy as the two halves of the scallop shell decelerate (close), then to reinvest this in acceleration as the shell is subsequently opened. This is simple for a pendulum, but it is no trivial matter for a machine that produces a high power output. When a kangaroo hops, it can store (in its tendons) about half of the energy with which it lands, for re-bounding. Cheng *et al.*² have shown that a scallop does considerably better — it recaptures almost all of the energy that ends the shell-closing phase for acceleration through the water in the next cycle.



The trajectory of a scallop such as *Placopecten magellanicus* during a short bout of swimming. Downward ejection of water from the gape provides the initial upward tilt, after which the jets and a slight inclination provide the thrust and lift that maintain its course. The animal claps at about 4 Hz, opens to about 13°, is tilted front-up at 6°–12° and moves at about 6 body-lengths per second. The scallop here is drawn several times larger than life relative to its course, and the inset, a view from above, shows the location of the jets.

This means that the resistance to acceleration (inertia) that the shell offers is not important; and the practical advantage of its thin, scalloped form must be in its lower gravitational mass, which reduces the need for lift.

Because they were aware that little energy is lost in oscillatory motion, and they knew the fluid-mechanical forces, Cheng *et al.*² were able to estimate the peak power output of the scallop adductor muscle *in vivo*. They came up with a value of 185 W kg⁻¹, which is remarkably high, and comparable to insect flight muscles or the best vertebrate striated muscles. Similarly, the maximum stress or force relative to cross-sectional area, of about 100 N m⁻², compares favourably with the best elsewhere. One implication of these data is that the scallop adductor operates under conditions that make especially good use of muscle as an engine. Moreover, the resilience of the scallop's elastic material, abductin, is over 90 per cent — better than the abductin of other bivalves, and surpassed only by the resilin found in insect wing hinges and other places in arthropods.

Unfortunately, the scallop is less im-

pressive in the lift and thrust that it gets from its efforts — that is, in the ratio of lift to drag of its hydrofoil, and in its overall (Froude) propulsion efficiency (the ratio of power output to power input). For a given area, long wings have better lift-to-drag ratios than short ones; and scallops, being almost circular, are short. Because they are denser than sea water, scallops also need lift, and the paired jets contribute little to this. A scallop whose half-shells are mirror images of each other (some species have nearly flat lower shells) gets its lift by moving with the plane of the shell tilted slightly above its path (see figure). But the circular design extracts its price, and the lift-to-drag ratios of scallops are closer to those of flying squirrels than to those of aircraft.

Propulsion efficiency is usually low in

jetting animals. For example, a squid uses several times as much power as does a fish of a similar size and speed that is propelled by its tail fin. The problem is that thrust is produced most efficiently by giving as large a mass of water as possible the smallest incremental speed — just the opposite of what a jet does³. And scallops are not an exception, although according to Cheng and DeMont¹ they rank highly among the jetters, mainly because they expel a large volume of water per clap.

But neither the lift-to-drag ratio nor the propulsion efficiency is likely to matter much to the survival of a scallop. Because they are aquatic animals scallops need relatively little lift, so minimizing drag is probably more important. They neither migrate nor chase prey, so propulsion efficiency may be less relevant than the production of thrust relative to engine size. Moreover, the lower operating costs of fin-based swimming may be offset by

1. Cheng, J.-Y. & DeMont, M. E. *Can. J. Zool.* **74**, 1734–1748 (1996).
2. Cheng, J.-Y., Davison, I. G. & DeMont, M. E. *J. Exp. Biol.* **199**, 1931–1946 (1996).
3. Vogel, S. *Life in Moving Fluids* 2nd edn (Princeton Univ. Press, 1994).

high construction costs and by other functional compromises. Although jetting incurs higher operating costs, it uses muscle in a simple and effective way.

A swimming bivalve is in no sense analogous to Samuel Johnson's dog that walks on its hind legs, nor (closer to home) to swimming by humans. Elegance isn't equivalent to biological effectiveness. Ironically, much of the recent work on scallop swimming has been reported in a journal of the government of Canada. Two

years ago, an international dispute followed the seizure by Canada of two US scallop boats for collecting on the continental shelf beyond the 200-mile limit. Its jurisdiction there covers only organisms that remain in physical contact with the seabed — Canada does not view scallops as swimmers. □

Steven Vogel is in the Department of Zoology, Duke University, Box 90325, Durham, North Carolina 27708-0325, USA.

RADIOASTRONOMY

A well-fed black hole

Mitchell C. Begelman and Joss Bland-Hawthorn

NGC1068 is a spiral galaxy in which large-scale gravitational disturbances are funnelling matter into the nucleus. A jet flowing from it is visible out to distances of several thousand light years, but the nucleus itself is hidden behind a screen of gas and dust. Radio observations have now pierced that screen, however, and they suggest that it conceals a black hole of 10 to 20 million solar masses. Surprisingly, gas is swirling into the hole at such a high rate, compared with the hole's mass, that the nucleus is radiating at close to the theoretical limit.

Most galaxies may have supermassive objects in their nuclei, but only a small fraction bear the hallmarks of a black hole that is being actively fed matter by the host galaxy. Such galaxies, which are said to possess an 'active galactic nucleus' (AGN), have very bright centres powered by a radiating accretion disk in rapid rotation about the black hole. The past few decades have seen a great deal of theoretical work on the structure of AGN, but it is only in recent years that observations have been good enough to restrain the theorists.

Late last year, a group of 50 astronomers gathered at Ringberg Castle in

the Bavarian Alps* to discuss probably the best observed of all active galaxies. NGC1068 is the archetype of a subset of active galaxies whose central engines are hidden from view — in the conventional picture, by a dense torus made of gas and dust. A direct view of the accretion disk would reveal an enormous flux of ionizing photons (from ultraviolet to gamma-ray wavelengths), but this radiation is absorbed or scattered by the dust and re-radiated at much lower energies, so most of it escapes in the infrared.

If all the infrared radiation we see is re-processed flux from the accretion disk, then the total luminosity could be as large as that of 100 billion Suns. Owing to the torus, this enormous energy output can only escape along a pair of opposite cones. They intersect the galactic disk to produce fans of ionized gas which can be seen out to many thousands of light years on both sides of the AGN — much like the beams of a lighthouse illuminating fog.

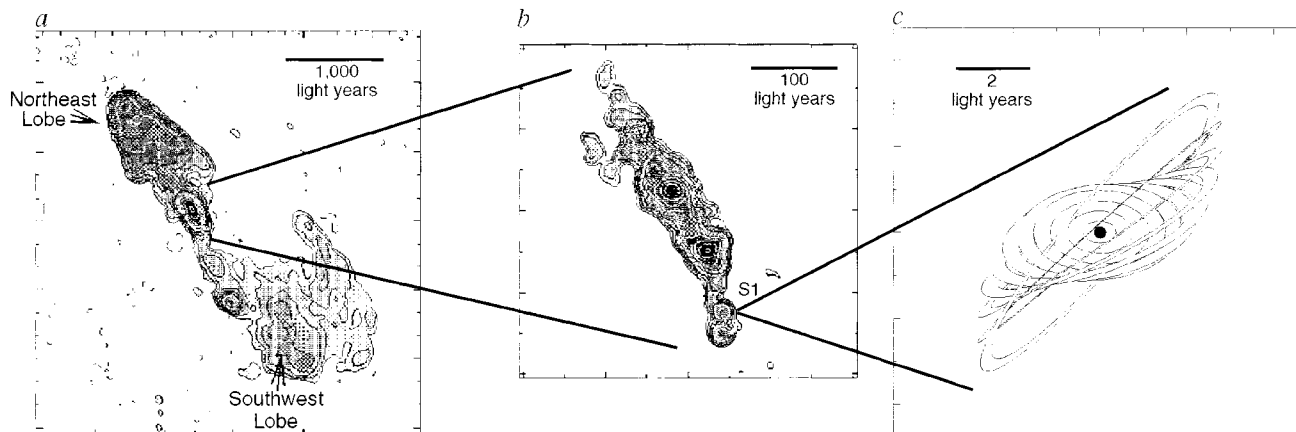
Molecular clouds within the ionizing cones produce strong mid-infrared radiation

but very little line emission from polycyclic aromatic hydrocarbons (PAHs). PAHs are grains less than ten ångströms across that are easily destroyed by X-ray radiation in terrestrial laboratories. The shape of the infrared spectrum and the weak PAH emission from the molecular clouds can be understood if the ionizing flux in these directions has a strong ultraviolet component (confirmed¹ by the infrared satellite ISO), and is several times more intense than the reprocessed flux observed along our line of sight. Some form of beaming seems to be necessary.

The biggest surprise comes from observations of water masers — small sites of intense, coherent microwave emission (L. J. Greenhill, Harvard-Smithsonian Center for Astrophysics; C. R. Gwinn, Univ. California at Santa Barbara). Their distribution throughout the S1 region (*b* in the figure) looks rather like the warped disk traced out by masers in NGC4258 (ref. 2), but with a much more pronounced warp. And their orbital velocities put the mass of the central black hole at about 10 to 20 million solar masses, implying that the combined luminosity at all wavelengths is about half the Eddington limit. That is the threshold at which the outward pressure of radiation overtakes the black hole's gravitational pull.

Radiation pressure must therefore have an important dynamical effect on the gas in NGC1068, at least close to the black hole, where gas and stars don't contribute much to the gravity. The expected directionality of radiation from a such a luminous accretion flow³ could conceivably account for the degree of intrinsic beaming required by analysis of molecular clouds within the ionizing cones.

But the strange geometry of NGC1068's nucleus resists explanation. Outside about 300 light years, interstellar gas seems to flow in an orderly, planar fashion through the galactic disk. Why, then, should the inner accretion disk be so severely misaligned with the galaxy that



Zooming in on NGC1068. The radio emission has two lobes at large scales^o (*a*) which are fed by inner jets^o (*b*). A supermassive black hole is hidden within the source marked S1. A thin, warped accretion disk

(*c*), illustrated with a series of discrete annuli, could be blocking our view of the black hole (black dot at centre). This innermost region is the source of the energetic particles that produce the radio emission.

both the axis of the ionization cone and the jet lie close to the galactic plane? Such misalignments are common in active galaxies, but nowhere else are they so clearly mapped.

Perhaps the fuel captured into the inner disk consists of giant molecular clouds whose orbits have been randomized by gas-dynamical or gravitational effects. One cloud of 100,000 solar masses could fuel activity in NGC1068 for 100,000 years — then the present alignment could reflect the motion of just one such cloud. But, unless individual clouds have very large masses, a disk a few light years across would need to be made of many clouds to maintain an adequate accretion rate, and the misalignment should then be much less pronounced. One possible explanation is a newly discovered instability which can cause an accretion disk to warp spontaneously under the action of radiation pressure from the compact source at its centre⁴ (c in the figure). It is a huge effect and can even cause the warp to wrap around on itself.

Other problems abound. Does the obscuring material really have the shape of a torus? Dynamical models have long faced the difficulty of finding a mechanism that will support such a torus against gravitational collapse to its orbital plane. The new evidence that NGC1068 is radiating at close to its Eddington limit makes it more plausible that radiation pressure acting on dust could support the torus against gravity. But the distribution of maser spots on the sky does not resemble what would be expected from a torus.

Moreover, if the maser emission is to be pumped by X-rays from the nucleus, as seems likely, the molecules must have a clear line of sight to the X-ray-emitting region. Such a line of sight is provided naturally by a warped disk, which at the same time could obscure our view of the nucleus and restrict the solid angle of the escaping ionizing radiation. Severely warped but geometrically thin accretion disks⁵ no longer seem as far-fetched as they once did. The new observations and diagnostic techniques presented at this workshop may force us to make such a radical shift in a long-held theory. □

Mitchell C. Begelman is at JILA, University of Colorado, Boulder, Colorado 80309-0440, USA. Joss Bland-Hawthorn is at the Anglo-Australian Observatory, PO Box 296, Epping, New South Wales 2121, Australia.

Axons follow Reelin routes

Anirvan Ghosh

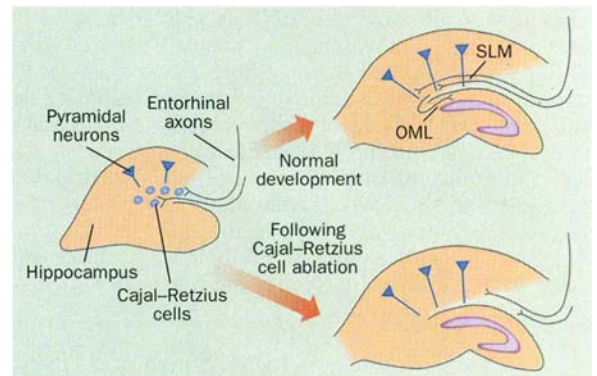
ONE of the most remarkable features of the developing nervous system is the ability of axons to identify and innervate distant targets. Not only must the axon grow in a particular direction, but it must stop growing when it reaches the appropriate target, before beginning the critical phase of synapse formation. In the mammalian central nervous system this problem is particularly challenging because the targets are often found several millimetres away, and literally millions of axons must find their targets within restricted developmental periods.

In a study published on page 70, Del Río *et al.*¹ provide evidence that in the developing hippocampus this problem may be solved, in part, by transient neurons that act as guideposts to direct axons to their final destinations. The study also shows that Reelin — a protein that has been implicated in the control of cell migration — may be involved in this guidance mechanism. Together with studies on the role of transient neurons in establishing connections in the cortex of the brain, these findings support an emerging principle that the developing brain contains populations of transient neurons that do not participate in mature functional circuits, but instead seem to act as a template for the formation of appropriate connections.

To address the mechanisms involved in the establishment of hippocampal connections, Del Río *et al.* studied connections between the entorhinal cortex and the hippocampus, that are thought to be involved in the formation of memories. They reconstituted the entorhinohippocampal pathway (EHP) by culturing explants of the entorhinal cortex with slices of the hippocampus. Entorhinohippocampal axons invaded the hippocampus and terminated in the appropriate layers — the stratum lacunosum-moleculare (SLM) and the outer molecular layer (OML; see figure).

During development, the SLM and the OML contain a transient population of neurons called the Cajal–Retzius cells, which are among the first cells to be generated in the hippocampus^{2,3}. To determine whether these cells are involved in directing innervation of the SLM and the OML, the authors asked whether the development of the entorhinohippocampal

axons would be affected if the Cajal–Retzius cells were ablated by treatment with specific agonists or toxins. They found that when explants from the entorhinal cortex were cultured with hippocampal slices from which the Cajal–Retzius cells had been ablated, the EHP connections did not form properly. Although the axons invaded the hippocampal tissue, only a few of them successfully innervated the SLM and the OML. Moreover, the axons that did innervate these layers were poorly developed and lacked the charac-



Summary of Cajal–Retzius cell ablation experiments in the hippocampus, performed by Del Río *et al.*¹. Ablation of the Cajal–Retzius cells leads to reduced innervation of the hippocampus, and this is thought to be due to the absence of the Reelin protein, which normally acts as a guidepost for axon development. SLM, stratum lacunosum-moleculare; OML, outer molecular layer.

teristic axonal branches, indicating that the Cajal–Retzius cells provide a signal that is necessary for proper development of the EHP.

To begin to identify the molecular cues that Cajal–Retzius cells may use, Del Río *et al.* examined the consequences of perturbing Reelin, which is normally expressed by these cells. The *reelin* gene was identified in 1995, and it is responsible for the *reeler* mutant mouse phenotype which is characterized by abnormalities in cell migration in the cortex and cerebellum^{4,5}. The authors used anti-Reelin antibodies to inhibit the function of the protein, and they found that development of the EHP was affected. Although the effects were

1. Lutz, D. *et al.* *Astrophys. Space Sci.* (in the press).
2. Miyoshi, M. *et al.* *Nature* **373**, 127–129 (1995).
3. Sikora, M. *Mon. Not. R. Astron. Soc.* **196**, 257–268 (1981).
4. Pringle, J. E. *Mon. Not. R. Astron. Soc.* **281**, 357–361 (1996).
5. Phinney, S. *Astrophys. J.* **347**, 29–51 (1989).
6. Gallimore, J. F., Baum, S. A., O'Dea, C. P. & Pedlar, A. *Astrophys. J.* **458**, 136–148 (1996).
7. Muxlow, T. W. B. *et al.* *Mon. Not. R. Astron. Soc.* **278**, 854–860 (1996).

1. Del Río, J. A. *et al.* *Nature* **385**, 70–74 (1997).
2. Super, H. & Soriano, E. *J. Comp. Neurol.* **344**, 101–120 (1994).
3. Soriano, E., Del Río, J. A., Martínez, A. & Super, H. *J. Comp. Neurol.* **342**, 571–595 (1994).
4. D'Arcangelo, G. *et al.* *Nature* **374**, 719–723 (1995).
5. Ogawa, M. *et al.* *Neuron* **14**, 899–912 (1995).
6. Ghosh, A., Antonini, A., McConnell, S. K. & Shatz, C. J. *Nature* **347**, 179–181 (1990).
7. Ghosh, A. & Shatz, C. J. *Development* **117**, 1031–1047 (1993).
8. Mackay, D. J. G., Nobes, C. D. & Hall, A. *Trends Neurosci.* **18**, 496–501 (1995).
9. Tessier-Lavigne, M. & Goodman, C. J. *Science* **274**, 1123–1133 (1996).

not as dramatic as those observed when the Cajal–Retzius cells were ablated, innervation of the SLM and the OML was markedly reduced, and the axons again failed to become branched. Moreover, analysis of the *reeler* mutant mouse showed a similar defect in formation of the EHP.

It is tempting to conclude that Reelin may act as a signal that specifies the patterning of the EHP, but a degree of caution is warranted. First, the effects of ablating the Cajal–Retzius cells are much more dramatic than the effects of the *reeler* mutation, suggesting that the consequences of functionally removing these cells cannot be entirely due to the loss of Reelin. Second, the defect in formation of the EHP in the *reeler* mouse seems to be temporary — although the EHP is less developed on the day of birth, by postnatal day 12, innervation by the entorhino-hippocampal axons is indistinguishable from controls. In fact, the *reeler* mutant mouse can be interpreted as having a defect in branching of EHP axons within the SLM and the OML, rather than in hippocampal innervation. Nevertheless, there is good evidence that development of the EHP does require Reelin, and that

this function is distinct from the part that Reelin plays in cell migration (movement of the cell body).

The involvement of hippocampal Cajal–Retzius cells in formation of the EHP is reminiscent of the way in which subplate neurons form a template for connections in the cerebral cortex^{6,7}. Normally, axons from the thalamus relay sensory information to the cortex through the thalamocortical projection. During development, as thalamic axons grow towards their cortical targets, they interact closely with the subplate neurons that line their path. When these neurons are ablated by injecting kainic acid into the developing cortex, the thalamic axons fail to innervate their appropriate cortical targets. So interactions between the thalamic axons and the subplate neurons are required for proper development of the thalamocortical projection. Moreover, the subplate neurons belong to a population of cortical neurons that are generated at an early stage of development and do not participate in the functional circuits of the mature cortex. The similarities between this system and development of the EHP indicate that interactions involving transient neuronal populations may be gener-

ally required for the patterning of connections in the mammalian brain.

Given that Reelin directs the development of hippocampal connections¹, might it also be involved in the patterning of thalamocortical connections? This seems to be unlikely as subplate neurons do not express Reelin^{4,5}, and thalamic axons can innervate the appropriate cortical targets in the *reeler* mouse. But this does not rule out the possibility that Reelin helps to regulate some other aspect of cortical connectivity. In the developing cortex, Reelin is expressed by cortical Cajal–Retzius cells known as the marginal-zone cells, which are generated at the same early stage of development as the subplate neurons. The marginal zone subsequently develops into a rich synaptic zone containing the dendrites of cortical pyramidal cells as well as axons from cortical and thalamic neurons. So the patterning of connections within the marginal zone could require the marginal-zone cells from the early stages of development, and it would be interesting to find out whether Reelin is involved in the development of these connections.

So Reelin — a protein that was initially associated with cell migration — now also seems to regulate aspects of axonal growth and guidance. This is particularly interesting in view of studies indicating that cell migration and growth-cone movement probably involve common intracellular mechanisms⁸. The experiments described by Del Río *et al.*¹ may be the first evidence that cell migration and axon guidance not only share intracellular molecular mechanisms, but that they might even be regulated by the same extracellular cues. It may not be too far-fetched to suggest that recently identified axonal-guidance cues, such as netrins and semaphorins⁹, could also regulate aspects of cell migration. In that case, two of the central aspects of nervous-system development, cell migration and axon guidance, could, in fact, be cellular manifestations of a common set of molecular interactions. □

Anirvan Ghosh is in the Department of Neuroscience, The Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21212, USA.

Corrections

● In “The sharp end of quasars” by Patrick S. Osmer (*Nature* 384, 415–416; 1996), in the sentence, “They were able to identify all the radio sources either as galaxies or as stars, and they conclude that none are quasars with $z > 5$ ”, the word “stars” should read “quasars”.

● In “A suspicious signature” by Adriano Aguzzi and Charles Weissman (*Nature* 383, 666–667; 1996), the source of the personal communication relating to the ‘target-cell hypothesis’ was Rudolf Karl Meyer, not K. H. Meyer as stated.

Spot the referee
IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS
IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Action-Plus Photographic

This completely unscientific contest was announced in *Nature* 384, 214 (1996), and drew on an old practice in British newspapers, the spot-the-ball competition. It had as its immediate inspiration a paper in the *British Medical Journal*, in which editors of the journal *Psychological Medicine* described a survey of the ability (or as it turned out, the lack of it) of authors of papers to identify the referees of the paper from their anonymous reports.

The spot-the-referee challenge for *Nature* readers was to identify the position, in x, y coordinates, in millimetres, of the referee in a picture of a soccer match from which his image had been removed. The specific target was the referee's nose. Here he is shown back in the action, and the official measure

of the position of his nose is $x, 40.5; y, 28$. Entries came from 18 countries. But all four of those who came closest, with estimates 0.5 mm out on the x axis and exact on the y , are in the United States — James Binley (Aaron Diamond AIDS Research Center, New York), Roberto Fernandez-Larsson (Children's National Medical Center, Washington DC), Scott A. Loiler (University of Massachusetts Medical Center, Worcester, Massachusetts) and Jijun Wan (University of California, San Diego). As in sport, but happily not in science, the referee's decision in this competition is final. Congratulations to the four winners (who will receive a year's subscription to *Nature*), commiserations to those who came close, and thanks to all who entered. □

Igneous ferment at Hamersley

Cornelis Klein

BANDED iron formations are iron-rich sedimentary rocks that were laid down under the oceans during the Archaean and Proterozoic aeons of Earth history (Fig. 1). They have a great deal to tell us about conditions in these earlier times, for instance helping us to assess the state of oxygenation of the atmosphere and oceans. They have also been taken as evidence that, especially during the period of maximum deposition 2,600–2,000 million years (Myr) ago, their depositional basins experienced a relatively quiet period with no major tectonic or magmatic activity on the sea floor. That conclusion comes under challenge with the appearance of the paper by Barley and colleagues on page 55 of this issue¹.

The most abundant and extensive banded iron formations (BIFs) occur in Western Australia, South Africa, Brazil and Canada, and are of Palaeoproterozoic age — that is, close to the Archaean–Proterozoic boundary^{2,3}. The best documented of them are those in the Hamersley Basin (Western Australia) and the Kaapvaal Craton (South Africa). Both of these huge chemical iron-formation deposits have generally been described as devoid of any close association, or substantial interlayering, of the BIFs with igneous rocks. So the source of at least the three most abundant oxide components in BIFs — SiO_2 , FeO and Fe_2O_3 — is generally thought to have been hydrothermal vents in the deep-ocean basins, the oxides having been transported to continental shelves where they were deposited.

Barley and colleagues¹, however, suggest that the Hamersley BIF sequence originated in different circumstances. The authors provide evidence for the existence of a large igneous province during the chemical precipitation of the Hamersley sequence, and conclude that the Hamersley BIF formed during a large tectono-magmatic event. Their seemingly insignificant Fig. 2a (page 56) is, in fact, a highly dramatic depiction of the relationship of BIF deposition to underlying magmatic activity.

After dating primary zircons in a thick rhyolite sequence above the BIF, and detrital zircons in an underlying sandstone, the authors have lumped all the igneous rocks — including dolerite, basalt and rhyolite — in the stratigraphic column of the Hamersley Group into one large igneous province. It was first proposed that volcanic activity was responsible for Hamersley BIF deposition over a quarter of a century ago⁴, but the direct association of a BIF immediately overlying basalt (as depicted in Barley and colleagues' Fig. 2a) is completely new. The Hamersley BIFs contain extensive stilpnomelane-rich meso- and microbands which have been interpreted as altered volcanic materials⁵; similar bands in the South African BIF sequences are thought to stem from Plinian (explosive) eruptions, with the very fine grain-sizes of the ash particles having resulted from far-away volcanic vents⁶.

The average chemical composition of BIFs throughout most of Precambrian time (not including the Fe_2O_3 -rich examples in the Neoproterozoic) shows SiO_2 usually ranging between 43 and 56 weight per cent; FeO between 17 and 30 wt%; Fe_2O_3 between 3 and 30 wt%; and CaO and MgO between 2 and 9, and 3 and 7 wt%, respectively². This relative constancy in average bulk composition over most of Precambrian time is commonly contrasted⁷ with the high variability in the stratigraphic sequences in which such BIFs occur. In Archaean stratigraphic columns containing BIFs, volcanic rock types are often present. But in some Proterozoic sequences the association of BIFs with possible igneous activity is less clear-cut⁸, even though some igneous associations are found in many.

The patterns of rare-earth elements in many

BIFs throughout the Precambrian (except, again, those of Neoproterozoic age) show very similar geochemical signatures; furthermore, these patterns are directly comparable to those from modern deep marine hydrothermal deposits^{2,3}. Therefore, even if there is no direct association of BIFs and underlying or coeval basalts

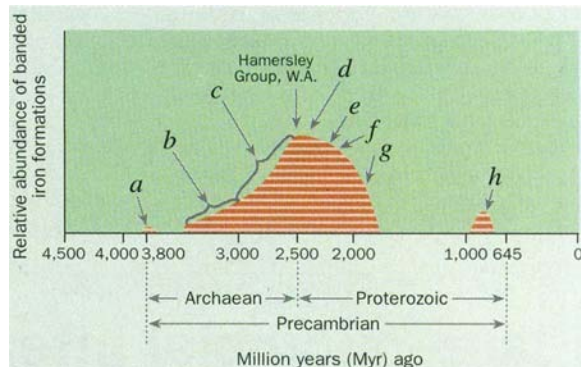


FIG. 1 The relative abundance of banded iron formations in the Precambrian. Estimated values are relative to those of the Hamersley Province, which is the largest BIF province in the world^{2,3}. a, Isua (West Greenland); b, Zimbabwe, South Africa, Ukraine, Venezuela, Western Australia; c, Canadian greenstone belts, Yilgarn block (Western Australia); d, Transvaal Supergroup (South Africa); e, Lake Superior region (USA); f, Krivoy Rog series (Russia); g, Labrador Trough (Canada); h, Rapitan Group (Canada), Urucum region (Brazil), Damara Supergroup (Namibia).

(or other rocks of igneous origin), it is pretty much generally accepted that the ultimate source of Si , Fe^{2+} and Fe^{3+} is a deep-ocean, seawater-dominated hydrothermal system.

So what is to be made of Barley and colleagues' view that the occurrence of a tectono-magmatic event during Hamersley time is contrary to the generally held idea of tectonic quiescence during BIF deposition? Somehow, in fact, the two geological processes must have coexisted. The Hamersley BIFs are the best worldwide example of three very well-developed scales of banding: macro-, meso- and microbanding. Of this, the microbanding (alternating bands of fine-grained quartz and an iron mineral on a 0.2–2-mm scale) is interpreted as primary depositional laminae resulting from annual seasonal variations in



FIG. 2 Aged rock bands — Fortescue Falls, Dales Gorge, part of one of the main Hamersley banded iron formations.

1. Barley, M. E., Pickard, A. L. & Sylvester, P. J. *Nature* **385**, 55–58 (1997).
2. Klein, C. & Beukes, N. J. in *Proterozoic Crustal Evolution* (ed. Condie, K. C.) 383–418 (Elsevier, Amsterdam, 1992).
3. Klein, C. & Beukes, N. J. in *The Proterozoic Biosphere: a Multidisciplinary Study* (eds Schopf, J. W. & Klein, C.) 139–151 (Cambridge Univ. Press, 1992).
4. Trendall, A. F. & Blockley, J. G. *Bull. Geol. Surv. W. Aust.* **119** (1970).
5. LaBerge, G. L. *Econ. Geol.* **61**, 147–161 (1966).
6. Beukes, N. J. & Klein, C. *Precamb. Res.* **47**, 99–139 (1990).
7. Gole, M. & Klein, C. *J. Geol.* **89**, 169–183 (1981).
8. Gross, G. A. *Can. Geol. Surv. Econ. Geol. Rep.* **22** (1965).
9. Trendall, A. F. *Geol. Soc. Aust.* **19**, 287–311 (1972).
10. Kasting, J. F., Holland, H. D. & Kump, L. R. in *The Proterozoic Biosphere: A Multidisciplinary Study* (eds Schopf, J. W. & Klein, C.) 159–163 (Cambridge Univ. Press, 1992).

basinal chemistry; in other words, they are chemical varves⁹. These beautifully laminated BIFs overlie the evidence of volcanic activity. Clearly, a very quiescent sedimentary basinal environment coexisted with a possibly very active magmatic period.

About 1,800 Myr ago, BIFs disappeared relatively abruptly from the Precambrian rock record. This is usually taken to be the consequence of oxidation of the deep-ocean waters which, as a result, became depleted in iron¹⁰. The rise in oxygen content from reducing conditions in the early Archaean to more oxidizing conditions in the late Palaeoproterozoic is thought to have been a gradual process.

At 2,470 Myr old, the BIFs of the

Hamersley Basin are part of that story, and they reflect the largest chemical sedimentary banded-iron pile in the world. Even if their Si and Fe components were derived directly from massive igneous activity, their sedimentary basinal conditions remained in harmony with the atmospheric and oceanic environment prevailing at that time. Nonetheless, the paper by Barley *et al.* will have a considerable impact on evaluation of the origin of BIFs, as well as on the significance attached to their unique occurrence in Precambrian time. □

Cornelis Klein is in the Department of Earth and Planetary Sciences, University of New Mexico, Albuquerque, New Mexico 87131, USA.

NITROGEN CYCLE

Plant-microbial interactions

Valerie T. Eviner and F. Stuart Chapin III

NITROGEN is an essential element for plant growth, and is available in various forms in soils: dissolved organic nitrogen (DON), ammonium (NH_4^+) and its nitrification product nitrate (NO_3^-). Two groups, writing in this issue^{1,2}, have examined nitrogen use by conifers and microbes in conifer soils. Taken together, their results imply that production and uptake of NO_3^- are more important in conifer ecosystems than previously appreciated, and that soil microbes — rather than plants — play the main biotic role in NO_3^- retention in these systems.

Kronzucker *et al.*¹ (page 59) show that white spruce, an important commercial species, is an NH_4^+ specialist; it has a limited capacity to use NO_3^- , which perhaps contributes to the frequent failure of its seedlings to compete with other plant species in disturbed sites. Stark and Hart² (page 61) demonstrate that soil microbes in conifer soils have a high capacity to both produce and immobilize NO_3^- , the form of soil nitrogen that is most vulnerable to loss from ecosystems.

It is generally assumed that NH_4^+ is the only form of inorganic nitrogen produced in large quantities in conifer systems because soil NO_3^- concentrations and net nitrification rates are typically quite low³. But Stark and Hart² show that high rates of autotrophic NO_3^- production occur in a wide

range of conifer soils, indicating that pools of soil NO_3^- are small because of rapid immobilization by microbes, not because of low rates of nitrate production. These results are surprising, because nitrifying microorganisms are often thought to be poor competitors for NH_4^+ compared with other soil microbes. The observations also contribute to growing evidence that autotrophic nitrification does occur in acid soils⁴.

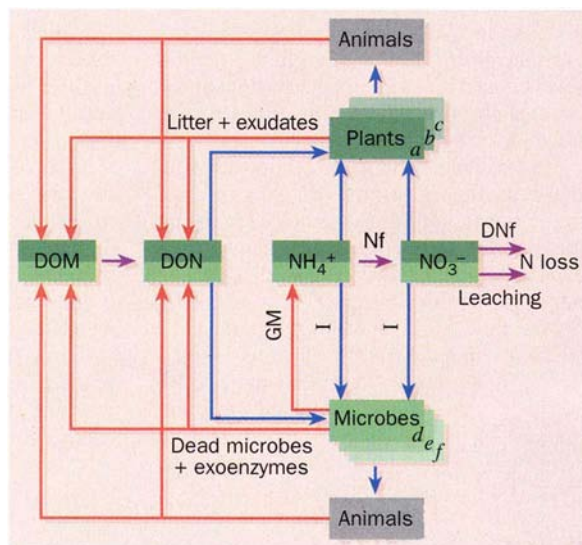


FIG. 1 A simplified, integrated view of the terrestrial nitrogen cycle. All nitrogen pools and transformations are influenced by both plants¹ (dark green) and soil microbes² (light green). Differences in nitrogen preference or retention among plant species (a, b, c ...) or among microbial groups (d, e, f ...) could shrink or swell any of the nitrogen transfers, with effects that propagate throughout the cycle. Blue arrows indicate uptake by organisms; red arrows inputs from organisms. DOM, dead organic matter; DON, dissolved organic nitrogen (such as amino acids and urea) that can be transported across membranes; GM, gross mineralization; I, gross immobilization; Nf, gross nitrification; Dnf, denitrification. This scheme ignores nitrogen deposition and fixation, pathways of nitrogen loss from the ecosystem, and fluxes between plant-microbial symbionts.

However, quantifying microbial uptake and transformations in the absence of plants, as Stark and Hart have done with their soil incubations, may not reflect processes occurring in the intact ecosystem. If white spruce¹ or its mycorrhizae⁵ are as effective in NH_4^+ absorption as laboratory studies suggest, uptake by conifers could limit NH_4^+ availability to nitrifiers; this effect has been observed in microcosm studies⁶. Equally, the strong preference for NH_4^+ shown by white spruce in Kronzucker and colleagues' solution cultures might not apply in the presence of soil microbes, which would compete with plants and their mycorrhizal fungi for NH_4^+ (ref. 6). Moreover, these authors show that white spruce did take up NO_3^- in the laboratory, when NH_4^+ was absent, with maximum induction after three days. The high nitrification rates reported by Stark and Hart could allow continuous induction of NO_3^- uptake by white spruce. Conversely, plants can alter the balance of inorganic nitrogen pools through differential uptake of NH_4^+ and NO_3^- (ref. 6).

Clearly, the study of soil microbes or plants in isolation is essential to document the physiological potential of organisms. But, in the intact ecosystem, realization of that potential depends on a wide variety of interactions and variables, as depicted in Fig. 1.

First there is the variety of forms in which soluble nitrogen is available. Most plants and microbes can absorb and use DON, NH_4^+ and NO_3^- , but species differ strikingly in their relative preferences, and those preferences depend on which of the forms of nitrogen are available^{1,7-9}. Given similar availability, energetic considerations must then be taken into account, and both plants and microbes should prefer DON (amino acids, for instance) over NH_4^+ over NO_3^- , because nitrate must be reduced to NH_4^+ , which must be added to a carbon skeleton to form amino acids, before newly assimilated nitrogen can be used for biosynthesis. Plants should be less restricted than microbes in using NO_3^- because they have their own source of energy for NO_3^- reduction, although some microbes derive sufficient energy for NO_3^- assimilation from litter².

On the supply side, nitrogen availability is defined by the pathway of transformations it undergoes. Assuming a simple pathway of dead organic matter (DOM) to DON to NH_4^+ to NO_3^- , with some uptake and conversion to plant or microbial biomass at each step, the relative supply rates must decrease in the order $\text{DON} \geq \text{NH}_4^+ \geq \text{NO}_3^-$. The decline in supply rates of these three forms will be greatest in nitrogen-limited ecosystems, where both plants and microbes absorb much of the nitrogen that becomes available at each step in the transformation chain.

Death of plant parts, microbes and animals, and excretion and exudation by

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

White spruce country in British Columbia, Canada. Here, however, aspen have outcompeted replanted conifers after clear-cutting. Kronzucker and colleagues' explanation¹ for this phenomenon is the conifers' comparative inability to use nitrate, as opposed to ammonium, as a nitrogen source.

these organisms, contributes to DOM and DON, reinitiating the cycle and leading to the same patterns of relative supply rate of different forms of soluble nitrogen. So even though the concentration of DON is quite low in most terrestrial ecosystems, its supply rate always exceeds that of NH_4^+ and NO_3^- , if nitrogen dynamics are dominated by internal recycling of nitrogen. This, of course, is not always the case — there can, for instance, be substantial inflow of inorganic nitrogen into rivers, and some terrestrial ecosystems are subject to large amounts of anthropogenically generated NH_4^+ and NO_3^- .

Preferential consumption of one form of nitrogen by a species modifies the balance of forms available to other species. For example, in ecosystems in which microbial growth and activity are carbon-limited, microbes use DON as a source of both carbon and nitrogen. They then excrete excess NH_4^+ , so that the supply rate of NH_4^+ approaches that of DON. This NH_4^+ is used by plants or nitrifiers, reducing its availability to either group more than would be expected from study of a single group in isolation. So knowledge of both relative supply rate and preferential consumption of specific nitrogen forms by individual species is essential to define the relative availability of different nitrogen forms to other species.

Other factors to consider are plant roots, mycorrhizae and the soil medium itself. Roots hold an energetic purse string that can shift competitive balances for different forms of nitrogen; so, with their more predictable energy supply, plants may be better able to reduce and assimilate nitrate than microbes, although not invariably so^{1,2,10}. Mycorrhizae involve a direct carbon transfer from the plant to its fungal symbiont, often allowing plants access to organically bound nitrogen that is not directly available to roots³. Root exudation and turnover also supply carbon to microbes in the rhizosphere, the soil environment influenced by individual roots, altering both production and consumption

of NH_4^+ and NO_3^- through changes in nitrogen mineralization and immobilization, and protozoan grazing of microbes.

Finally, the heterogeneity of the soil medium strongly influences the interactions among organisms and, consequently, net ecosystem fluxes. Most rhizosphere soils have higher rates of nitrogen transformations and show different net fluxes from bulk soils. At a smaller scale, microsite heterogeneity in carbon and nitrogen supply influences the balance between immobilization and

production by different plant and microbial groups^{2,6}, and therefore the forms of nitrogen available to individual organisms. Ion mobility influences the spatial extent of localized hot spots and depletion zones, with anions such as NO_3^- diffusing more rapidly than NH_4^+ along concentration gradients. At the scale of individual cells, species preference can reflect kinetics of influx and efflux¹, or proximity to sites of production (an example is the absorption of DON by decomposer and mycorrhizal fungi following exoenzymatic breakdown of dead organic matter).

As must now be evident, the challenges in understanding the terrestrial nitrogen cycle remain considerable. For example, as commonly measured, microbial biomass lumps together quite distinct groups — decomposer microbes that can compete with plants for nitrogen, mycorrhizal fungi that funnel nitrogen to plants, nitrifiers that use the product of nitrogen mineralization as their own energy source, and soil microfauna which graze on roots and microbes. The new studies^{1,2} are important in showing that different organisms have radically different capacities to process various forms of nitrogen and carbon. The next step — no easy one — will be to put all of these actors together on the same stage.

Valerie T. Eviner and F. Stuart Chapin III are in the Department of Integrative Biology, University of California at Berkeley, Berkeley, California 94720-3140, USA.

1. Kronzucker, H. J., Siddiqi, M. Y. & Glass, A. D. M. *Nature* **385**, 59–61 (1997).
2. Stark, J. M. & Hart, S. C. *Nature* **385**, 61–64 (1997).
3. Vitousek, P. M. et al. *Ecol. Monogr.* **52**, 155–177 (1982).
4. Laanbroek, H. J. & Woldendorp, J. W. *Adv. Microb. Ecol.* **14**, 275–304 (1995).
5. Read, D. J. *Experientia* **47**, 376–391 (1991).
6. Norton, J. M. & Firestone, M. K. *Soil Biol. Biochem.* **28**, 351–362 (1996).
7. Recous, S., Mary, B. & Faurie, G. *Soil Biol. Biochem.* **22**, 913–922 (1990).
8. Chapin, F. S. III, Moilanen, L. & Kielland, K. *Nature* **361**, 150–153 (1993).
9. Kielland, K. *Ecology* **75**, 2373–2383 (1994).
10. Davidson, E. A., Stark, J. M. & Firestone, M. K. *Ecology* **71**, 1968–1975 (1990).

Beam alignment

NYLON, polyester and a few other polymers can be orientated. Stretch a sample to a few times its original length, and its long-chain molecules all line up in the stretching direction. The resulting fibre or film has remarkable strength and tenacity in that direction — making it ideal for cord, tape, textiles, packaging, and similar duties. Daedalus now has a scheme to orientate a whole new set of polymers.

Most commercial polymers, of course, are made by coupling small monomer molecules together, either catalytically or by irradiation. Imagine, says Daedalus, a single energetic quantum or particle traversing a monomeric fluid. As in a cloud-chamber, it will leave a trail of free radicals or ionized molecular fragments. Polymeric chains will grow out from these sites in random directions. But if the trail is dense enough, and chain growth is terminated efficiently enough, the chains will just about manage to link up with each other before being quenched. The result will be a single, long, polymeric molecule, with a number of short side-chains, extending along the track of the initiating particle.

A dense enough particle beam should therefore convert the whole specimen into a perfectly aligned parallel-chain polymer. Both the length and direction of the orientated molecules would be determined by the beam characteristics.

A nuclear reactor or synchrotron seems the best source of collimated protons or X-rays; but with luck a laser system will be almost as effective. A whole new range of orientated plastics will result — tough, fibrous solids with a wood-like grain running through them. Even such intractable thermosetting polymers as Bakelite and alkyd resin will show a new, more flexible side to their natures. Like wood, these new structural materials could be broken down by pulping, and made into novel papers, fibres and textiles.

Furthermore, a second irradiation at right angles to the first could crosslink their parallel chains into molecular sheets. Like molecular chains, these would probably crumple and interleave in the melt or in solution, giving a quite new sort of polymeric solid, neither glass, rubber nor monolith. And if the two irradiations were conducted together, they might weave a sort of molecular cloth, with two sets of parallel fibres interdigitated at right angles. As with cloth cut on the bias, you could stretch this cunning product one way, when it would shrink in the other. So a sample cut too short could still be stretched to fit. Clumsy engineers everywhere would rejoice.

David Jones

Mary Leakey 1913–96

FEW individuals have both the opportunity and the ability to transform their chosen field. Mary Leakey was such a scientist — she combined a flair for fieldwork with a legendary capacity for hard work and a formidable talent for analysis. This conjunction of abilities enabled her to change our understanding of human evolutionary history, with discoveries that are unrivalled in their variety and significance. And although her early years in research were spent under the shadow of her husband, Louis Leakey, in later life she justly came to be regarded as a remarkable archaeologist in her own right.

Mary Leakey's career in prehistory began in the 1930s, when excavation techniques were crude and stone tools were looked upon more as trophies than as objects of scientific study. Her first digs were in France and England, but her practical skills and excavation technique were honed during her first African site investigation at Hyrax Hill in Kenya. Between 1943 and 1947, she and Louis worked at the Middle Pleistocene site called Olorgesailie, where she developed and refined the excavation methods that later allowed the spatial relationships between artefacts and animal bones to be recorded and preserved on the so-called 'living floors' at Olduvai Gorge.

As well as her contributions to the Miocene and Plio-Pleistocene hominid fossil record, Mary Leakey's discoveries included the earliest direct evidence of hominid activity — the remarkable trails of footprints at Laetoli in Tanzania — and painstaking recordings of East African cave paintings. But of all these achievements, her research at Olduvai Gorge stands out in terms of both the results and the methods that were pioneered there.

The name Olduvai — which literally means 'the place of the wild sisal' — was given by the Masai to a gorge that extends for some 50 km into the Serengeti plains on the western margin of the Eastern Rift Valley in Tanzania. The main valley and 'Side Gorge' were produced by the erosion of sediments that had been deposited in the beds of lakes and streams. Interleaved between the layers of sand and clay that make up the bulk of the sediments are layers of volcanic ash from eruptions of the (now extinct) volcano called Olmoti. The glass in these ash layers, combined with the well-preserved record of the Earth's magnetic field in the other sediments, provided the reliable dating framework that was to become an important component of

the scientific significance of Olduvai Gorge.

In 1935, Mary Leakey made her first visit to the place that she was to work in, and around, for over half a century. Until the mid-1950s, visits for survey and excavation were intensive, but sporadic. Many stone tools and mammal fossils were found, but there was no sign of the human ancestors who had so clearly

been active there. Then, in 1959, a hominid molar tooth was found at site MNK in Bed I. Several weeks later, when Mary Leakey was examining what they had concluded might be a 'living floor' associated with crudely fashioned 'Oldowan' stone tools, she found the first fragments of an early hominid cranium. It was given the designation OH 5 and was initially referred to *Zinjanthropus boisei*, but it is now more usually placed in *Paranthropus boisei*.

This discovery was a turning point because it enabled the Leakeys to attract much-needed funds to mount a sustained research programme at the gorge. Jack Evernden, a Berkeley geochemist, was invited to use the relatively new technique of potassium-argon dating to determine the absolute age of the sediments. With his colleague Garniss Curtis he proposed that OH 5 was just less than 2 million years old. By providing absolute ages for many of the other ash layers, Evernden and Curtis established the first absolute chronology for any early hominid site, and Olduvai Gorge soon became the yardstick by which the dates of other sites could be calibrated.

The early 1960s saw the discovery of the first of a series of hominid remains that were to be referred to *Homo habilis* — a hominid with a larger brain and smaller chewing teeth than *Paranthropus boisei*. Until 1964, publications announcing the discovery of hominid remains were all in the name of Louis Leakey. But thereafter, Mary Leakey's

important, and then dominant, role in the research was more justly reflected in the authorship. She announced new hominid finds, including a cranium, OH 24, of *Homo habilis* and the limb bones of *Homo erectus*, and proposed new interpretations of the hominid remains from Olduvai.

Mary Leakey's 1971 monograph on the archaeological evidence from Beds I and II secured the widespread acceptance of her identification and interpretation of the Oldowan and the Developed Oldowan cultures of Beds I and II. Similarly, a more recent monograph in the Olduvai Gorge series presented scrupulously documented interpretations of the archaeology from the later sediments. All of this work was done in close collaboration with Richard (Dick) Hay, and their scientific partnership — which continued when they switched their attention to Laetoli — has not always been given the credit it deserves.

The discovery of hominid fossils from considerably older sites elsewhere in East Africa rekindled Mary Leakey's interest in Laetoli. Here, in 1974 and 1975, her team recovered fine examples of *Australopithecus afarensis* and, in subsequent years, she excavated footprint trails including several made by hominids. At around 3.7 million years old, these hominid footprints are still the earliest direct evidence of hominid behaviour, and they show that hominids of this antiquity could walk upright.

Once an archaeologist has excavated a site, important components of the information it contains — such as the spatial relationships between the artefacts — are lost forever. Mary Leakey realized that archaeologists have a special responsibility to record not only what they judge to be important at the time, but also evidence for others coming after them. Her meticulous excavation and recording of the archaeological record at Olduvai Gorge is a unique resource for testing hypotheses about artefact function and hominid land use. That alone would be a rich legacy, but coupled with her other achievements, they add up to a truly remarkable contribution to the study of human evolutionary history.

Bernard Wood

Bernard Wood is in the Hominid Palaeontology Research Group, University of Liverpool, PO Box 147, Liverpool L69 3BX, UK.

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS
IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

John Reader/SPL

Biology's new Rosetta stone

SIR — Completion of the sequencing of the entire yeast (*Saccharomyces cerevisiae*) genome is a major event in the history of biology: all those involved are to be congratulated. Roughly 6,200 proteins from an estimated 4,000 homologous protein families are encoded on the 16 yeast chromosomes, undoubtedly representing most eukaryotic protein families. The analysis of these data will provide powerful tools

for understanding the genomes of other eukaryotes, as well as the more recently completed bacterial sequence, that of *Methanococcus jannaschii*¹. There is concern, however, that, in the competitive rush to sequence and analyse genomes, the required investment of time and money to organize and integrate this wealth of data will not be made.

Four groups have already begun initial

efforts to organize these data: the *Saccharomyces* genome database project at the Department of Genetics, Stanford University Medical School; the Martinsried Institute for Protein Sequences (MIPS); the GeneQuiz² consortium (involving the European Molecular Biology Laboratory, the CNB-UAM in Spain and the AIC-SRI in the United States); the European Bioinformatics Institute GeneCrunch³ resource and the Yeast protein database⁴.

These are independent efforts encompassing the following: a physical map of genes; their assignment of probable (biochemical) func-

tion; some pattern-match information from the program PROSITE⁵; and pointers to protein structures⁶. They do not, however, have a common nomenclature or a controlled vocabulary, and contain no direct references to various known relationships between yeast genes, such as those that are on the same metabolic or regulatory pathway.

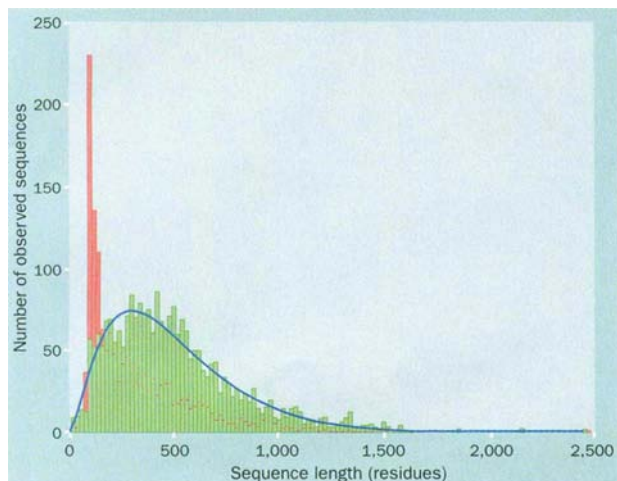
The distribution of yeast gene lengths in these databases suggests a high proportion of complex multidomain proteins and a systematic error in the assignment of open reading frames (ORFs). The figure plots the length distributions of the ORFs for which the GeneCrunch³ resource had 'clear' matches to earlier determined sequences and for those with no significant match. The peak at 100 to 110 amino acids, three times higher than at other lengths, is seen only in the unmatched distribution, and clearly indicates an artefact in the definition of potential ORFs.

It is disturbing that this problem was not noted in any of the initial data analyses. Although the MIPS and GeneCrunch resources are now aware of this anomaly, it still exists in the newly released *Bacillus subtilis* sequences (available at <http://bacillus.genome.ad.jp>) and among the many GenBank-listed mouse ORFs.

An examination of the yeast genome as presented in these resources suggests that nearly 60% of ORFs have been assigned a probable function, primarily by BLAST, FASTA or PROSITE software matches. We have confirmed most of these initial assignments, plus an additional few, using a more rigorous local dynamic programming⁷ comparison with SwissProt, a new functional pattern database⁸, and with *S. cerevisiae* itself. In our survey, if one requires confirmation of functional assignment by agreement between two different methods, only about 45% of the yeast ORFs have a cleanly assigned regional function.

As might be expected, the most common enzymatic function (by nearly a factor of two) is the transfer of phosphate groups within the metabolic, signalling and regulation pathways. Next are the hydrolases (acting on ester and peptide bonds), followed by the oxidoreductases, the ligases, lyases and the isomerases.

Caution must be applied to any interpretation of this function hierarchy. The proteins with assigned functions are those that have been most studied, and they may therefore be disproportionately represented in the sequence databases, making it easier to identify their yeast homologues. In addition, many of the functional assignments are the result of second-hand, inherited annotation. There are, for example, yeast genes that have been assigned tentative functional family membership, but with no information as to how the annota-



Sequence length distributions of *S. cerevisiae* proteins listed in GeneCrunch. ORFs that have 'clear' identifiable homologues are shown in green and those that do not match previously assigned proteins are shown in red. The probable artefact in the latter distribution contains more than 400 sequences. The former distribution seems to fit a χ^2 distribution with five degrees of freedom rather than, for example, a Poisson distribution, although we have no obvious explanation for this.

COMPARISON OF COMPLETED GENOMES

<i>S. cerevisiae</i>	Z-score	<i>M. jannaschii</i>
YPR118w (hypothetical)	49.138	MJ0454 (translation-initiation factor, eIF-2B, subunit α)
YPR118w (hypothetical)	27.740	MJ0122 (translation-initiation factor, eIF-2B)
YGL169w (SUA5 protein, translation-initiation protein)	23.262	MJ0062 (unknown function)
YNL040w (hypothetical)	25.772	MJ0564 (alanyl tRNA synthetase)
YKL215c (hypothetical)	91.577	MJ0963 (N-methylhydantoinase)
YPL175w (N-acetylglucosaminyltransferase)	47.874	MJ1178 (unknown function)
<i>B. subtilis</i>	Blast score	Matches
YABH_PBS (unknown function)	20	MJ1104 (homoserine kinase)
YABR_PBS (unknown function)	21	MJ0117 (translation-initiation factor)
	9	YJR007w (translation-initiation factor)
YAAD_PBS (unknown function)	111	YMR096w (hypothetical)
Similar product in <i>Haemophilus influenzae</i> , yeast and Archaeobacterium <i>Methanococcus vannielii</i>	111	YNL333w (hypothetical)
	111	YFL059w (hypothetical)
	126	MJ0677 (ethylene-inducible protein homologue)
	33	gi1574495 (hypothetical, in <i>H. influenzae</i>)
	34	gi297176 (expressed ORF, in <i>Stellaria longipes</i>)
	33	gi44737 (unidentified reading frame, in <i>M. vannielii</i>)

Top, Homologies between hypothetical proteins, or those of unknown function, with proteins of assigned biochemical function (given in parentheses) in *S. cerevisiae* and *M. jannaschii*. Six pairs are listed for which the similarity is better than 20 s.d. from the mean. The similarity statistics were calculated using the Smith-Waterman dynamic programming comparison⁸ of all 6,200 *S. cerevisiae* ORFs with all 1,738 *M. jannaschii* ORFs. Bottom, Hypothetical *B. subtilis* proteins that are homologous to proteins in *M. jannaschii* and/or *S. cerevisiae*. These were identified using BLASTp and confirmed by multi-alignment pattern analysis⁹. The blast scores are presented in equivalent numbers of aligned amino-acid identities. The YAAD *B. subtilis* protein, which is homologous to three yeast genes, one *M. jannaschii* and three others, all of unknown function, is highly conserved across the three kingdoms. It contains, for example, 37 shared amino-acid identities in a run of 40, along with many other conserved positions.

tion was originally assigned to the matched sequence. Finally, and more important, such assignments of homology are generally only at the biochemical level and provide very limited information about 'biological' function — similar biochemistry does not necessarily imply similar biology.

Caution must also be applied when considering the homologous functional assignments given to multidomain proteins. Among the yeast proteins longer than 600 amino acids, there are at least 170 with newly assigned functions where the annotation was inherited from a match to a sequence or pattern matching only a small fraction of yeast protein's length. Even when yeast genes are compared to other yeast genes, the potential exists for such confusion in defining homologous families.

A comparison of all yeast genes to themselves identified more than 500 ORFs that cannot simply be resolved into separate homologous clusters, together with about 1,800 that resolved into nearly 700 clusters defined by a common, similar amino-acid sequence. On average, the members of these clusters are randomly distributed over the 16 chromosomes, that is, the number of distinct chromosomes containing members from a particular cluster is proportional to cluster size. In addition, members of different clusters have similar neighbours on different chromosomes. These facts suggest that many transposon or other recombination events in the history of *S. cerevisiae* are possibly related to its domestication.

At least 1,867 of the yeast ORFs are preceded by two or more previously recognized yeast regulatory elements^{9,10}, excluding the TATA box. As described in the Transfac database, MIG1 is a yeast regulatory DNA-binding protein involved in glucose utilization. YBL086c is the only hypothetical protein matched in a set of 10 other genes by the gal-10 MIG1 binding site (R04145 in Transfac). The 14 genes originally used to define the MIG1 binding site additionally contain the site for the repressor of CAR1 expression. This site is also found in the upstream region of the hypothetical protein YBL086c, supporting its involvement in sugar metabolism.

Only six of the 6,200 promoter regions examined contain the long MAT α 1 sites. One of these, the upstream region of the hypothetical YBL028c gene, is additionally matched by an RAP1 binding site involved in repressing the silent mating loci. Surely this gene has a biological function associated with mating-type regulation in *S. cerevisiae*? Such analyses support the need to integrate fully our various yeast resources.

Sequence-comparison methods between the various recently sequenced genomes is a powerful tool for cross-species identification of function (table, top section), raising provocative new questions. In one case (table, bottom section), there is a very highly conserved gene product, of

unknown function, shared among the three representative genomes of the three terrestrial kingdoms, *M. jannaschii*, *B. subtilis* and *S. cerevisiae*. Given its conservation, its function must be rather fundamental. But because of the complexities of multiple-domain proteins, and the missing regulatory and metabolic database pointers, much work remains to be done to exploit fully such comparative analyses. It is to be hoped that a major collaborative effort will be undertaken to provide a fully integrated yeast genome database that will provide biology with a new 'Rosetta stone'.

Sudeshna Das, Lihua Yu

Chrysanthos Galtatzes, Robert Rogers

James Freeman, Jadwiga Blenkowska

R. Mark Adams, Temple F. Smith

BioMolecular Engineering Research Center,

College of Engineering,

Boston University,

Boston, Massachusetts 02215, USA

James Lindelien

Time Logic Inc.,

Incline Village, Nevada 89451, USA

1. Buit, C. J. et al. *Science* **273**, 1058–1073 (1996).
2. Scharf, M. et al. in *Proc. 2nd Int. Conf. Intelligent Systems Mol. Biol.* (eds Altman, R. et al.) 348–353 (AAAI, Menlo Park, 1994).
3. Schneider, R. et al. in *Proc. Supercomput. '96 Semin.* (Saur, Mannheim, in the press).
4. Garrels, J. I. *Nucleic Acids Res.* **24**, 46–49 (1995).
5. Bairoch, A. *Nucleic Acids Res.* **20**, 2013–2018 (1992).
6. Abola, E. E., Bernstein, F. C., Bryant, S. H., Koetzle, T. F. & Weng, J. in *Crystallographic Databases — Information Content, Software Systems, Scientific Applications* (eds Allen, F. H., Bergerhoff, G. & Sievers, R.) 107–132 (Int. Union Crystallog., Bonn, 1987).
7. Smith, T. F. & Waterman, M. S. *J. Mol. Biol.* **147**, 195–197 (1981).
8. Adams, R. M., Das, S. & Smith, T. F. *Protein Sci.* **5**, 1240–1249 (1996).
9. Ghosh, D. *Nucleic Acids Res.* **20**, (suppl.) 2091–2093 (1992).
10. Wingender, E., Dietze, P., Karas, H. & Knuppel, R. *Transcript. Factors* **24**, 238–241 (1996).

Induced synthesis of plant volatiles

SIR — Wounding of plant tissue by insect feeding triggers the release of volatile compounds, which have been implicated in the attraction of natural enemies of the insect herbivores^{1–8}, and are derived from several biosynthetic pathways. The release of indole and several terpenes is delayed, and is induced only by interaction of damaged plant tissues with a substance (or substances) in the oral secretions of insects^{9,10}. We have used ¹³C₂ to show that cotton plants (*Gossypium hirsutum* L.) damaged by feeding beet armyworms (*Spodoptera exigua* Hübner) synthesize these induced volatile compounds *de novo*, as distinct from the conversion of some previously prepared precursor. Furthermore, ¹³C-labelled volatiles are released from the leaves soon after synthesis and are rapidly replaced by unlabelled compounds when ¹³C₂ is replaced by ¹²C₂.

This is the first demonstration of *de novo* synthesis of volatiles induced by insect damage to a plant, showing the

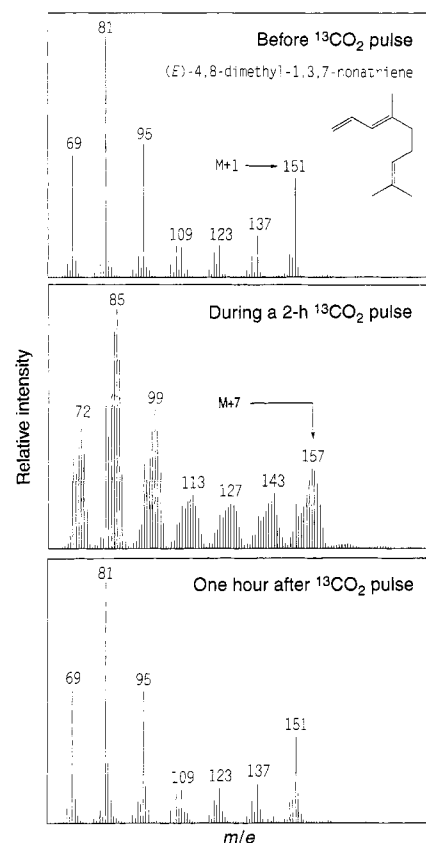


FIG. 1 Chemical ionization (isobutane) mass spectra of (E)-4,8-dimethyl-1,3,7-nonatriene released from beet armyworm-damaged cotton plants immediately before the introduction of ¹³C₂ in synthetic air (top), during a 3-h collection in which a 2-h pulse label with ¹³C₂ in synthetic air was administered (middle) and a 3-h collection starting 1 h after the pulse label ended (bottom). Laboratory-reared¹⁹, third instar beet armyworm larvae, starved for 7 h, were placed on six-week-old, greenhouse-grown cotton plants (variety Deltapine Acala 90) at 1500 h, and allowed to feed continuously. Filtered air (1 l min⁻¹) passed down over the plants contained within a 34 × 6-cm diameter glass sleeve and out of a hole in a split plate at the base that closed loosely around the stem of the plant²⁰. Synthetic premixed air (360 p.p.m. ($\geq 99\%$ ¹³C) CO₂, 20.7% O₂ and a balance of N₂) was introduced at 0900 h on day 2 by flushing the chamber at 10 l min⁻¹ for 2 min and then reducing the flow to 1 l min⁻¹ for 2 h. The same procedure was used to switch back to atmospheric air. Plant volatiles were collected between 0600 and 2100 h in 3-h intervals by pulling half of the air passing over the plant through Super Q adsorbent (25 mg) traps around the base of the volatile-collection chamber; the rest of the air vented out from the bottom of the system. Volatiles were eluted from the adsorbent traps with 150 μ l CH₂Cl₂; 400 ng each of *n*-octane and nonyl acetate were added as internal standards and 1- μ l aliquots analysed by gas chromatography-mass spectroscopy on a Finnigan MAT ITS40 (ion trap) mass spectrometer.

dynamic role of plants in interactions with the second (insect herbivores) and third (parasitoids and predators) trophic levels.

For several of the volatiles that are released, the delay in emission after herbivore attack suggests that chemical changes

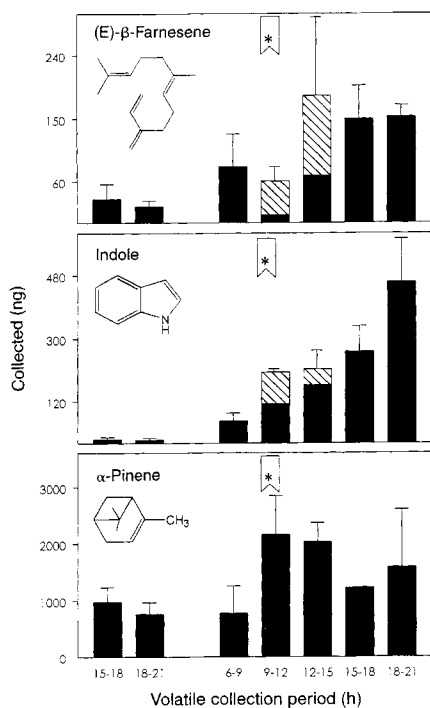


FIG. 2 Incorporation and turnover of ^{13}C label for (E)- β -farnesene, indole and α -pinene from beet armyworm-damaged cotton plants (hatched areas, ^{13}C -labelled product; solid areas, ^{12}C -labelled product; inverted arrow with asterisk, 2-h period during which $^{13}\text{CO}_2$ was administered). See Fig. 1 for details. Volatiles were analysed by capillary gas chromatography and quantified by flame ionization detection using the peak area of added nonyl acetate as an internal standard. Gas chromatography chemical ionization mass spectroscopy was used to calculate $^{13}\text{C}/^{12}\text{C}$ ratios²¹. Selected mass ions were quantified via computer software analysis. By summing the intensities of the $(M+2)^+$ through $(M+n)^+$ (n , number of carbon atoms) and comparing this with the intensity of the molecular ion signal, the contribution of ions associated with enriched and unenriched molecules and the $^{13}\text{C}/^{12}\text{C}$ ratio for each compound was determined. Mean ($\pm$ s.e.m.) volatile release is shown with each data point, based on three replicates.

are occurring in the plant. The lag between beet armyworm feeding damage to cotton leaves and the release of induced volatiles, is almost a day. The volatiles, most of which are terpenes, exhibit a diurnal cycle of emission with a great increase in release during the photophase, which coincides with the period when parasitic wasps are foraging for hosts⁹. In contrast, constitutive compounds are released immediately after physical damage to the plant, at a rate correlated with the amount of damage.

It has been suggested that one of the very common volatile terpenes induced by herbivore damage, the homoterpene (E)-4,8-dimethyl-1,3,7-nonatriene, is stored as a β -glucoside and is cleaved by the β -glucosidase contained in the regurgitant of the feeding insect¹¹. In fact, a β -glucosidase from *Pieris brassicae* (cabbage-white butterfly) larvae has been reported as an

elicitor of herbivore-induced cabbage-plant odour¹².

Determining whether volatiles are released as a consequence of cleavage of glucoside-bound moieties or as a result of *de novo* biosynthesis, and whether the mechanism that controls these processes differs in different plant species, is critical to understanding herbivore-induced plant-volatile release.

To assess the pace of synthesis and metabolic turnover, we exposed cotton plants to $^{13}\text{CO}_2$ at atmospheric concentrations for a 2-h or a 36-h period, so maintaining a steady-state condition during the incorporation of the label¹³. The mass-spectral data for (E)-4,8-dimethyl-1,3,7-nonatriene (Fig. 1), as well as for five other terpenes that were emitted by the plant in response to insect feeding, demonstrate that most of the molecules for each of these induced compounds had incorporated ^{13}C label. But only a small share of the molecules exhibited total incorporation of ^{13}C after the 2-h pulse-label treatment. All the volatiles identified are released by the caterpillar-damaged plants and not by the caterpillars themselves^{14,15}.

Within the isoprenoid pathway, insect feeding did not regulate the synthesis of compounds uniformly, as demonstrated by the high variability of ^{13}C incorporation into individual terpenes. In response to insect damage, the plant synthesized several of the acyclic terpenes including (E)- β -ocimene, linalool, (E,E)- α -farnesene, and (E)- β -farnesene (Fig. 2), as well as the homoterpenes (E)-4,8-dimethyl-1,3,7-nonatriene and (E,E)-4,8,12-trimethyl-1,3,7,11-tridecatetraene (which are formed by a series of degradative steps from sesquiterpene and diterpene precursors, respectively¹⁶). The rapid loss of carbon marker subsequent to the pulse label indicates a high metabolic turnover rate for these compounds.

In contrast, many of the cyclic isoprenoid products, including the monoterpenes α -pinene (Fig. 2) and β -pinene, as well as the sesquiterpenes caryophyllene and α -humulene, were not labelled during the 2-h exposure to $^{13}\text{CO}_2$, and showed very low levels of ^{13}C incorporation even with a 36-h exposure. The slow rate of turnover for these constitutive terpenes is in agreement with turnover for several terpene-accumulating plants, including garden sage and lodgepole pine, in undamaged conditions¹⁷. The monoterpenes limonene and myrcene incorporated low, but detectable, levels of ^{13}C label during the 2-h $^{13}\text{CO}_2$ exposure, and 20% and 35% ^{13}C incorporation, respectively, during the 36-h $^{13}\text{CO}_2$ pulse interval.

The biosynthesis of indole (Fig. 2) confirms that a second synthetically distinct and distant pathway is activated in response to insect feeding. Products of both the isoprenoid and the tryptophan

pathways are released within roughly the same time interval after plant damage from insect feeding. Indole, which seems to be a pivotal intermediate in the biosynthesis of the plant growth hormone indole acetic acid (IAA)¹⁸, turns over rapidly. Labelled indole precursors had passed through the pathway within 2 h after the $^{13}\text{CO}_2$ exposure was stopped.

These findings clearly show that, in cotton, the volatiles induced by insect herbivore damage are synthesized *de novo* with little or no release of these compounds from storage. The elicitor or elicitors from the feeding insect activate at least two different biosynthetic pathways. The synthesis of these metabolites in response to insect feeding unmistakably points to an active reaction that channels energy towards the release of volatiles as a defensive response to herbivore attack.

These biochemical data demonstrate that the plant is active in mediating the interactions between the second and third trophic levels. Indeed, the fact that the plant synthesizes the herbivore-induced volatiles *de novo* indicates that this specific response to herbivore damage is a reliable signal to natural enemies of herbivores as to the location of their hosts.

P. W. Paré

J. H. Tumlinson

Center for Medical, Agricultural, and Veterinary Entomology,
United States Department of Agriculture
Agricultural Research Service,
1700 Southwest 23rd Drive,
Gainesville, Florida 32608, USA
e-mail: tumlnsn@nervm.nerdc.ufl.edu

1. Turlings, T. C. J., Tumlinson, J. H. & Lewis, W. J. *Science* **250**, 1251–1253 (1990).
2. Dicke, M. J. *Plant Physiol.* **143**, 465–472 (1994).
3. R  se, U. S. R., Manukian, A., Heath, R. R. & Tumlinson, J. H. *Plant Physiol.* **111**, 487–495 (1996).
4. Dicke, M. et al. *J. Chem. Ecol.* **16**, 381–396 (1990).
5. Dicke, M., Sabelis, M. A., Takabayashi, J., Bruin, J. & Posthumus, M. A. *J. Chem. Ecol.* **16**, 3091–3118 (1990).
6. Turlings, T. C. J. et al. *Proc. Natl Acad. Sci. USA* **92**, 4169–4174 (1995).
7. Steinberg, S., Dicke, M. & Vet, L. E. M. *J. Chem. Ecol.* **19**, 47–60 (1993).
8. Mattiacci, L., Dicke, M. & Posthumus, M. A. *J. Chem. Ecol.* **20**, 2229–2247 (1994).
9. Loughrin, J. H. et al. *Proc. Natl Acad. Sci. USA* **91**, 11836–11840 (1994).
10. Par  , P. W. & Tumlinson, J. H. *Fl. Entomol.* **79**, 93–103 (1996).
11. Boland, W., Feng, Z., Donath, J. & G  bler, A. *Naturwissenschaften* **79**, 368–371 (1992).
12. Mattiacci, L., Dicke, M. & Posthumus, M. A. *Proc. Natl Acad. Sci. USA* **92**, 2036–2040 (1995).
13. Robinson, T. *Science* **184**, 430–435 (1974).
14. Turlings, T. C. J., Tumlinson, J. H., Eller, F. J. & Lewis, W. J. *Entomol. Exp. Appl.* **58**, 75–82 (1991).
15. Turlings, T. C. J. et al. *J. Chem. Ecol.* **17**, 2235–2251 (1991).
16. Donath, J. & Boland, G. W. *J. Plant Physiol.* **143**, 473–478 (1994).
17. Gershenzon, J., Murtagh, G. J. & Croteau, R. *Oecologia* **96**, 583–592 (1993).
18. Normann, J., Slovin, J. P. & Cohen, J. D. *Plant Physiol.* **107**, 323–329 (1995).
19. King, E. G. & Leppla, N. C. in *Advances and Challenges in Insect Rearing* (US Gov. Printing Service, Washington DC, 1984).
20. Heath, R. R. & Manukian, A. *J. Chem. Ecol.* **20**, 593–608 (1994).
21. Sch  rmann, W., Ziegler, H., Kotzias, D., Sch  nwitz, R. & Steinbrecher, R. *Naturwissenschaften* **80**, 276–278 (1993).

Y chromosomes of Jewish priests

SIR — According to biblical accounts, the Jewish priesthood was established about 3,300 years ago with the appointment of the first Israelite high priest. Designation of Jewish males to the priesthood continues to this day, and is determined by strict patrilineal descent. Accordingly, we sought and found clear differences in the frequency of Y-chromosome haplotypes between Jewish priests and their lay counterparts. Remarkably, the difference is observable in both the Ashkenazic and Sephardic populations, despite the geographical separation of the two communities.

The human Y chromosome has useful properties for studies of molecular evolution^{1,2}. Except for the pseudo-autosomal region, it is inherited paternally and does not recombine. It can be used to construct patrilineal genealogy cladograms comple-

than paternal descent by which male Jews are assigned to the priesthood. Identification as a priest carries with it certain social and religious obligations which have tended to preserve this identity within Jewish communities. Based on surveys of Jewish cemetery gravestones, priests represent approximately 5% of the estimated total male world Jewish population of roughly 7 million (data not shown).

We identified haplotypes of 188 unrelated Y chromosomes using the polymerase chain reaction (PCR) applied to genomic DNA isolated from buccal mucosal swab samples from Israeli, North American and British Jews. We constructed haplotypes using first, the presence or absence of the Y Alu polymorphic (YAP) insert, thought to represent a unique evolutionary event dated to between 29,000 and 340,000 years ago^{1,5}; and second, a

trast, we found no significant difference in the distribution of alleles for the non-Y-chromosome locus polymorphism D1S191 (data not shown). These Y-chromosome haplotype differences confirm a distinct paternal genealogy for Jewish priests.

We further identified subjects as being of Ashkenazic or Sephardic origin. This refers to the two chief, separate communities which developed within the diaspora during the past millennium⁹. As shown in the table, the same haplotype distinction can be made between priests and lay members within each population. This result is consistent with an origin for the Jewish priesthood antedating the division of world Jewry into Ashkenazic and Sephardic communities, and is of particular interest in view of the pronounced genetic diversity displayed between the two communities⁹. This conclusion is further supported by the relative preponderance of the YAP⁺, DYS19B haplotype in both populations, suggesting that this may have been the founding modal haplotype of the Jewish priesthood.

Taken together, our findings define a set of Y chromosomes of recent common origin. Differences which have accumulated in the genomic DNA of the Y chromosomes of Jewish priests during the relatively short time since the establishment of the priesthood, should be useful in defining rates and mechanisms of Y-chromosome evolution.

Karl Skorecki

Sara Selig

Shraga Blazer

Bruce Rappaport Faculty of Medicine and Research Institute and Rambam Medical Centre, Technion-Israel Institute of Technology, Haifa 31096, Israel

and Department of Medicine, University of Toronto,

Toronto M5S1A8, Canada

email: skorecki@tx.technion.ac.il

Robert Bradman

Neil Bradman

P. J. Waburton

Monica Ismajlowicz

Department of Biology, University College London, London WC1E 6BT, UK

Michael F. Hammer

Laboratory of Molecular Systematics and Evolution,

Biosciences West, University of Arizona, Tucson, Arizona 85721, USA

1. Hammer, M. F. *Nature* **378**, 376–378 (1995).
2. Jobling, M. A. & Tyler-Smith, C. *Trends Genet.* **11**, 449–456 (1995).
3. Benerecetti, A. S. S. et al. *Genetic Diversity Among Jews* (eds Bonne-Tamir, A. & Adam, A.) 45–50 (Oxford Univ. Press, New York, 1992).
4. Ritte, U. et al. *J. Mol. Evol.* **37**, 435–440 (1993).
5. Hammer, M. & Horai, S. *Am. J. Hum. Genet.* **56**, 951–962 (1995).
6. Santos, F. R. et al. *Hum. Genet.* **97**, 309–313 (1996).
7. Roewer, L. et al. *Hum. Mol. Genet.* **5**, 1029–1033 (1996).
8. Gyapay, G. et al. *Nature Genet.* **7**, 246–249 (1994).
9. Motulsky, A. G. *Nature Genet.* **9**, 99–101 (1995).

Alleles	HAPLOTYPE FREQUENCY <i>F</i> (standard error)					
	All		Ashkenazic		Sephardic	
	Cohen <i>n</i> =68	Israelite <i>n</i> =120	Cohen <i>n</i> =44	Israelite <i>n</i> =81	Cohen <i>n</i> =24	Israelite <i>n</i> =39
YAP [−] DYS19 A	0.162 (0.045)	0.091 (0.026)	0.205 (0.061)	0.074 (0.029)	0.083 (0.056)	0.129 (0.054)
B	0.544 (0.060)	0.325 (0.042)	0.454 (0.075)	0.321 (0.052)	0.709 (0.093)	0.333 (0.075)
C	0.162 (0.045)	0.300 (0.042)	0.227 (0.063)	0.272 (0.049)	0.042 (0.041)	0.359 (0.077)
D	0.088 (0.035)	0.083 (0.024)	0.091 (0.044)	0.111 (0.035)	0.083 (0.056)	0.026 (0.024)
E	0.029 (0.020)	0.017 (0.012)	0.000 —	0.025 (0.017)	0.083 (0.056)	0.000 —
YAP ⁺ DYS19 (all)	0.015 (0.014)	0.184 (0.035)	0.023 (0.024)	0.197 (0.045)	0.000 —	0.153 (0.057)
<i>P</i> χ^2	<0.001		<0.01		<0.01	

Ashkenazic, Jewish communities of northern Europe; Sephardic, Jewish communities of north Africa and the Middle East; Cohen, Priest; Israelite, lay Jew. A–E, different DYS19 haplotypes.

mentary to those formulated using maternally inherited mitochondrial DNA.

The phenotypic differences that exist between different communities of contemporary Jews in the world are thought to emanate, at least in part, from genetic admixture with neighbouring communities of non-Jews, during a prolonged dispersion^{3,4}. The genetic basis of this diversity has been investigated using analysis of neutral DNA markers, including mitochondrial and Y-chromosome markers⁴. However, previous studies have not considered the subsets of male Jews comprising the priesthood (Cohanim). Significantly, there is no procedure other

polymorphic GATA repeat microsatellite, DYS19 (refs 6, 7). We also typed a subset of samples for the non-Y-chromosome CA-repeat polymorphism, D1S191 (ref. 8).

We determined the designation of each subject as a member of the priesthood by direct questioning. Subjects who were not sure of their designation or who identified themselves as 'Levite' (a separate junior priesthood, based on a different, less-well-defined patrilineal lineage) were not included in the current analysis.

We identified six haplotypes, whose frequencies are shown in the table (YAP[−] DYS19A–E and YAP⁺ DYS19, all alleles). Applying the χ^2 test to the frequencies of the Y-chromosome haplotypes distinguishes priests from the lay population. The most striking difference was in the frequency of YAP⁺ chromosomes among priests compared with lay Jews. Only 1.5% of Y chromosomes among priests were YAP⁺, in comparison to a frequency of 18.4% in lay Jews. In con-

Scientific Correspondence

Scientific Correspondence is intended to provide a forum in which readers may raise points of a scientific character. Priority will be given to letters of fewer than 500 words and ten references.

Artificial laboratories

Michael Berry

Would-be Worlds: How Simulation is Changing the Frontiers of Science. By John L. Casti. Wiley: 1996. Pp. 242. £17.99, \$24.95.

In his *Discourse on Method*, published in 1637, Descartes described how he understood the essential physics of rainbows by carrying out what we would now call a numerical simulation. At the frontier of physics in those days was Snell's law of refraction, which Descartes used to plot sun-rays through a raindrop. He discovered that the emergent rays were concentrated at an angle near 42° from the direction back towards the Sun, and thereby explained the rainbow as directional focusing (an 'angular caustic'). Recently, I have been using Maxwell's laws of physics, together with a small computer and modern mathematical software, to calculate the absorption of microwaves in a sphere, as part of a study of possible effects of radar on human eyes. Both models are sophisticated and crude at the same time, albeit in different ways. Descartes ignored diffraction and polarization, whereas I include it — on the other hand, a homogeneous dielectric sphere far more faithfully represents a raindrop than an eye.

Philosophically, there seems little difference between what Descartes did three-and-a-half centuries ago and what I and many other physical scientists are doing today. The cultural difference is,

however, enormous. Descartes' simulation was a substitute for an analytical solution that he was unable to find. But small computers are now so powerful, so easy to program, and so readily available that simulation has become routine in situations where there is no hope of encapsulating the phenomenon under investigation in a simple formula, and where the governing laws are known imprecisely, if at all. I have often thought that this change in scientific practice deserves serious and systematic study, so I approached this book with high expectations.

The author surveys the application of computers for simulating our world — and various imaginary ones — in a wide range of situations, including numerical weather forecasting, traffic flow in Albuquerque, New Mexico, and the evolution of artificial life inside a computer. Along the way, he discusses many subtle and important questions, such as when it is desirable for a model to be an accurate representation of external reality, when such accuracy might actually hinder the acquisition of insight and how complexity can emerge as large-scale behaviour in a system of many components that individually behave simply.

I do have some criticisms. The book is disorganized, in the sense that many of the examples in chapters that ostensibly deal with different topics ("Pictures as programs", "The science of surprise", "Artificial worlds", "The reality of the virtual") could have been interchanged without the reader noticing. And recent research on the stability of the Solar System does not fit naturally in the lists of "surprise generators" or "nature's unanswered or unanswerable questions".

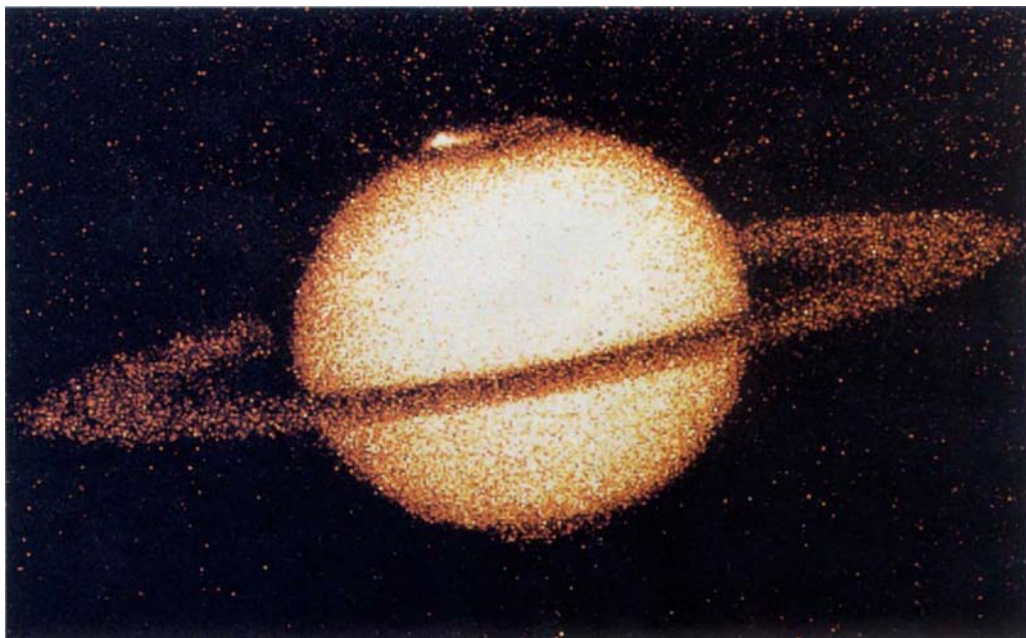
I was disappointed not to see any mention of quantum computers, whose different logic provides hope for the solution of problems that are uncomputable in the traditional way. A further criticism is that the colour plates are poorly reproduced and, in most cases, add little to their black-and-white duplicates in the text — in the traffic simulations, colour is important but almost impossible to see.

On the whole, the writing is lively and straightforward, although readers will wince at the phrase "it's place" and smile at the description of "nondeductive modes of reasoning — induction and abduction".

As a scholarly treatment of its subject, Casti's book fails. But as a popular account of the enormous change computers have brought to the everyday practice of science (Descartes notwithstanding), it is an easy and enjoyable read, and I can happily recommend it.

Michael Berry is at the H. H. Wills Physics Laboratory, University of Bristol, Tyndall Avenue, Bristol BS8 1TL, UK.

An account of the Hubble Space Telescope's troubled history and its importance in advancing our understanding of space introduces the photographs and clear explanations in *Hubble's Universe: A New Picture of Space* by Simon Goodwin. In this Hubble ultraviolet image of Saturn, the aurora at the north pole is clearly visible. Constable, £14.95.



Travels of a neuro-romantic

Chris McManus

The Island of the Colour-blind. By Oliver Sacks. Picador/Knopf: 1996. Pp. 345. £16.99, \$24.

ISLANDS are romantic and mysterious. Sacks identifies many of the obvious sources for such feelings, including *The Island of Dr Moreau*, Darwin and Wallace, Robert Louis Stevenson, Gauguin, the statues of Easter Island, the mutiny on the *Bounty*, and *Lord of the Flies*. Like most of us, Sacks could not resist the temptation of what he describes as "impulsive and unsystematic trips" to the Pacific. Not that unsystematic, though, for he made a veritable Cook's tour for neurologists.

He visited Pingelap and Pohnpei in Micronesia where *maskun*, an autosomal recessive colour-blindness (achromatopsia), affects eight per cent of people. He also went to Guam, where for 40 years the cause of the enigmatic *lytico-bodig* — a neurodegenerative disorder reminiscent of Parkinson's, motor neuron

disease and Alzheimer's — has defied identification; and it must be said that the problem is hardly elucidated further in Sacks's book.

His stated motivations for his travels are straightforward: "not part of any programme or agenda, not intended to prove or disprove any thesis, but simply to observe". Simply to observe? If we believe Popper, then that simply cannot be done. "I went," writes Sacks, "as a neurologist, or neuroanthropologist, intent on seeing how individuals and communities responded to unusual endemic conditions." One imagines a Malinowski or Pitt-Rivers, braving heat and disease, trekking through unmarked terrain and learning languages never written down. But this is actually "from the major television series", and we are disabused of any remaining romance on reading how the party arrived on Pingelap equipped not only with medical and film equipment but even with bottled water.

How successful is *The Island of the Colour-blind* as 'neuroanthropology'? Not very. With its rambling, over-romanticized style (one is tempted to say 'highly coloured'), it consists mostly of vast digressions, frequently concerning Sacks's hobby since childhood of botany, all supple-

mented by lengthy endnotes comprising a quarter of the book.

If the book has a core it has to be the colour-blindness of its title. The topic has a venerable history. In his *Remarks on Colour*, Wittgenstein asked us to "Imagine a tribe of colour-blind people, and there could easily be one". Pingelap and Pohnpei are as close as one will get, with congenital achromatopsia far more common than the typical one in 30,000. A devastating typhoon in 1775 left Pohnpei with just 20 people. The population quintupled in the next decades, and in 1820 the first of two colour-blind children was born to the *Nahnmwarki*, the leader of the island. Now eight per cent of the population is affected, and half the population carries the gene.

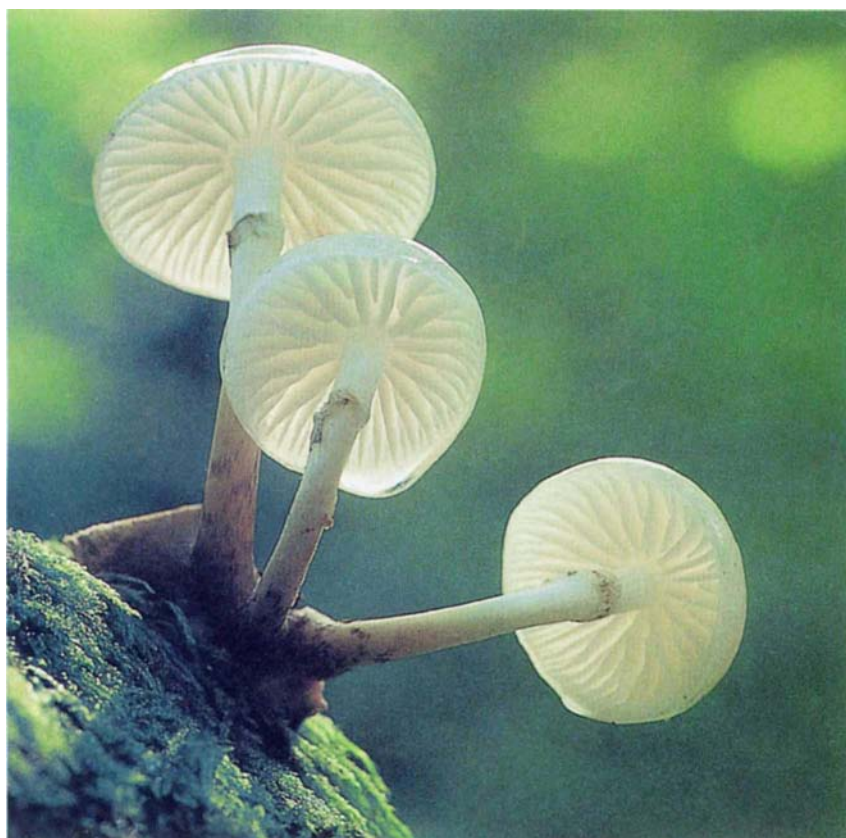
Achromatopsia is perhaps a misleading name. An absence of retinal cones, responsible for colour vision in daylight, restricts vision to that carried out by rods, which work properly only for scotopic, night-time, vision. This means that normal daylight seems blindingly intense. Absent foveal cones also disrupt fixation and produce severe nystagmus — an involuntary movement of the eye. Colour-blindness in itself is in some ways the least of the problems of those affected. It is nonetheless psychologically and philosophically the most interesting.

Needless to say, achromatopsics have no concept of colour and no real understanding of it, although they recognize that 'normals' use colour words that are learned for use metaphorically. Does Sacks go much beyond this? Not really. Ultimately he is a neuro-romantic, toying with ideas of compensation and enhanced abilities, feeling sure that achromatopsics must surely have hyperacuity for form, texture or whatever, only on the basis of anecdotal evidence. He even suggests they have "very acute auditory and tactual memories" — which would be a bit like a person with one leg necessarily having a better sense of smell.

Mysteriously, Sacks is surprised that achromatopsics can see stars at night as their "acuity is a tenth of normal". Photopic acuity, yes, but their scotopic acuity should be unaffected, and it seems to be. Sometimes a little theory is useful.

His analysis of the sociocultural implications comes to little, and one realizes he is more neurologist than anthropologist. Eventually he asks: "Was [this] an island of the colour-blind after all?" 'No' has to be the answer. Most of the population had normal vision and rapidly recognized and sometimes ostracized those who did not. Wittgenstein would not have wanted to visit Pohnpei after all. □

Chris McManus is at Imperial College School of Medicine at St Mary's, London W2 1PG and in the Department of Psychology, University College London, Gower Street, London WC1E 6BT, UK.



THE Roman poet Horace wrote: "Meadow mushrooms are in kind the best; it is ill trusting any of the rest." This is still sound advice for any fungi fancier unable to identify species found in the wild. *The Wild Mushroom* by George McCarthy comments on which species are good to eat but it is no identification guide — rather, a photographic celebration of the beauty of the subject. Fountain, £24.95.

Current affairs

George Philander

Currents of Change: El Niño's Impact on Climate and Society. By Michael H. Glantz. Cambridge University Press: 1996. Pp. 194. £40, \$59.95 (hbk); £14.95, \$19.95 (pbk).

HENRY David Thoreau asserted that "men are never tired of hearing how far the wind carried men and women, but are bored if you give them a scientific account of it". If he is correct, then Michael H. Glantz has written an absorbing book. Its emphasis is on detailed descriptions of how and when El Niño has brought floods, droughts and other disasters to various parts of the globe, not on scientific explanations.

Glantz brings to the fascinating and important topic of El Niño the perspective of a social scientist, so the book is mainly concerned with the impact that this series of climatic changes has on society. Physical science is not Glantz's area and his book only hints (sometimes in a confusing manner) at the reasons why El Niño is scientifically exciting and challenging. Glantz should nonetheless be commended for his effort to bridge the gulf between scientists and those who benefit from science.

Many people have heard of El Niño because of articles in magazines as diverse as *Reader's Digest* and *Soybean Digest*, but few have more than a vague notion that it is 'something' happening in the waters 'somewhere' off the coast of Peru and Ecuador. In this well-produced book, which has many informative tables and figures, they will find much information about El Niño, the recurrent natural phenomenon that influences weather and climate globally.

The first part of the book recounts "emerging interest" in the subject. Initially scientists studied, as unrelated phenomena, the oceanic aspects of El Niño — the appearance of unusually warm surface waters off the coast of Peru and Ecuador where guano production and fisheries are adversely affected — and its atmospheric aspects which include failure of the Indian monsoon. Only in the late 1950s, when data describing sea-surface temperature patterns over the entire tropical Pacific first became available, did scientists recognize that these continually changing patterns are not only created by the atmosphere but also affect it profoundly; El Niño is a product of interactions between the tropical Pacific Ocean and the atmosphere.

Only after the devastating El Niño of 1982–83 did scientists put their research to practical use and begin activities that are leading to the routine prediction of El Niño. Such forecasts are of value



Mexico's answer to the Domesday Book is the *Relaciones Geográficas*, a sixteenth-century survey of the country carried out by the Spanish colonialists. Barbara E. Mundy describes this feat as well as looking at the Amerindian contributions to mapmaking in *The Mapping of New Spain: Indigenous Cartography and the Maps of the Relaciones Geográficas*. University of Chicago Press, \$40, £31.95.

to everyone because El Niño affects everyone, either directly by a change in climatic conditions or indirectly because of the global nature of the economy. El Niño can decrease the harvest of anchovies off Peru and of coconuts in the Philippines, so affecting the price of commodities ranging from soaps and detergents with coconut oil as an ingredient, to fishmeal fed to chickens, and to soya beans that can provide a substitute for fishmeal. Commodities traders and brokerage firms are certainly aware that El Niño affects them but the public at large, and many government officials whose decisions affect agriculture, public health and water resources in many countries, apparently do not appreciate that El Niño predictions amount to 'usable' science.

Glantz points out that "forecasts of El Niño now abound" and is concerned about "too much emphasis... on developing an El Niño forecast capability and, at the same time, an underemphasis on the value of using existing El Niño information". He nonetheless points out that predictions of El Niño, including the most recent episodes, have been erratic. The

discussion of how predictions are made is poor, the section on computer models especially so, and there is no mention of what the future is likely to bring by way of improved forecasts.

Scientists have some skill in predicting climate changes that are common to all El Niño episodes. With fairly simple models, they anticipate variations in the sea-surface temperature of the eastern equatorial Pacific averaged by area. Forecasts for other parts of the globe are then possible on the basis of past records that indicate statistically significant correlations between that parameter and climate parameters in other regions. To do better requires the approach of weather forecasters — prediction of how spatially varying fields that describe temperature, rainfall and so on will evolve over a certain period. A report for laymen on the exciting efforts being made to start such operational climate forecasts remains to be written. □

George Philander is in the Department of Geosciences, Princeton University, Guyot Hall, Princeton, New Jersey 08544-1003, USA.

The sky isn't falling!

Peter Ward

Cretaceous-Tertiary Mass Extinctions: Biotic and Environmental Changes.

Edited by Norman MacLeod and Gerta Keller. Norton: 1996. Pp. 575. \$55, £40. **Dinosaur Extinction and the End of an Era: What the Fossils Say.** By J. David Archibald. Columbia University Press: 1996. Pp. 237. \$57, £40 (hbk); \$29, £20 (pbk).

The Great Dinosaur Extinction Controversy. By Charles Officer and Jake Page. Addison-Wesley: 1996. Pp. 209. \$25.

RESEARCH in many disciplines of the geological sciences was transformed following the appearance in 1980 of a now celebrated paper by Luis Alvarez and colleagues. Dealing with the mass extinction that ended the 'age of dinosaurs', the paper generated a minor extinction pulse of its own — in trees killed for printing all the subsequent discussion of the topic. Now, as we approach the paper's twentieth anniversary, three new books join the mountain of published work that has already stemmed from it.

What many people are not aware of is that the Alvarez paper offered not one hypothesis, but two. First, the authors suggested that the Earth was hit by a large comet or asteroid some 65 million years ago. Second, they argued that the result was a mass extinction that wiped out many Mesozoic species including, most famously, the dinosaurs on land, and ammonites in the sea.

That Earth was hit 65 million years ago by a large body about 10 km across was confirmed to most scientists' satisfaction within five years of Alvarez's original publication. The physical scars of such an impact from numerous Cretaceous-Tertiary boundary sites around the globe, together with the later discovery of a candidate crater (Chicxulub crater in the Yucatan peninsula in Mexico), have provided a preponderance of evidence that the Alvarez team was indeed correct in its first premise. That this impact was the main cause of a mass extinction, however, has been far more difficult to prove — or disprove. The best we have is a temporal coincidence of the impact with an extinction. This may be as good as the evidence will ever get; it has certainly been sufficient to convince many Earth scientists working on the problem.

Yet sceptics remain and, if the enormous amount of palaeontological research so far has not convinced them, perhaps nothing ever will. We are now seemingly in a phase analogous to a stubborn military retreat, with the pro-impact

extinction group moving inexorably forward, and the anti-impact forces slowly giving up territory ("there may have been impact-generated extinction in the tropics, but not at higher latitudes"). But they are fighting a grim rearguard action as they retreat, taxonomic group by taxonomic group, biotic province by biotic province.

That some palaeontologists still remain sceptical is clear from these three very different books about the controversy. The multiauthored tome edited by Norman MacLeod and Gerta Keller is aimed at specialists, and results from a conference held in Boston several years ago. Its 20 chapters bring together a large mass of work, yielding much new interesting information. It spans the taxonomic and geographic gamut of Cretaceous-Tertiary locations, organisms and sedimentary patterns characteristic of the end of the Cretaceous period. Most chapters deal either with the stratigraphic distribution of Cretaceous organisms or with sedimentology or sea-level changes. Many of the authors of the chapters on sea-level change have difficulty divorcing observation from interpretation. Half of the chapters have as coauthor at least one leading critic of the impact hypothesis, such as Keller, MacLeod or Stinnesbeck. There are some useful chapters dealing with hypothesis testing, especially those by MacLeod and David Archibald, which I found first rate.

Archibald has expanded his chapter to produce a truly excellent book, *Dinosaur Extinction and the End of an Era*. Archibald and I are at opposite ends of the extinction controversy, but I highly recommend his book, especially to a general readership. It is lucid, well written, and presents the data on vertebrate, and especially dinosaur, extinction in a forthright and lucid way.

For the most part, these two books accept as fact that the Earth was indeed hit 65 million years ago. But *The Great Dinosaur Extinction Controversy*, by Charles Officer and Jake Page, discards as heresy even this largely accepted tenet. To begin with, one might be forgiven for thinking that this book was written as a parody, especially when one comes across sentences in the preface such as: "Indeed, most of the 'science' performed by the Alvarez camp has been so inexplicably weak and the responses to it so eagerly accepted by important segments of the scientific press (never mind the popular press and the tabloids) that some skeptics have wondered if the entire affair was not, on the impact side, some kind of scam." Unfortunately, one soon realizes that this is a grim rearguard action — something akin to the Roman Catholic Church's view of Bruno and Galileo in the seventeenth century.

There are some wonderful new views of the exciting years immediately following Alvarez's publication, as seen by the 'loyal opposition'. Yet there is much distortion

in this book, such as the use of my data and publications (while never mentioning my own interpretation of the data) supposedly to "prove" that the ammonites became extinct gradually, not suddenly — just the opposite of my findings.

Judging from these three books, it is clear that the dinosaurs, although long dead, are keeping the nature of their vanishing still controversial.

Peter Ward is in the Department of Geological Sciences, University of Washington, Seattle, Washington 98195, USA.

New in paperback

Designing the Molecular World:

Chemistry at the Frontier by Philip Ball. Princeton University Press, \$16.95, £12.95. "The whole book is pervaded by a sense of authority and comprehension", Peter Atkins, *Nature* **358**, 550 (1994).

At Home in the Universe: The Search for the Laws of Complexity

by Stuart Kauffman. Penguin, £8.99. "This is a courageous book... I guarantee that any reader whose imagination has survived an academic education — or has never been exposed to one — will learn a lot, and be changed forever", Ian Stewart, *Nature* **379**, 33 (1996).

The Faber Book of Science

edited by John Carey. Faber, £9.99. "Everything an anthology should be: it is entertaining, stimulating and occasionally startling. Carey's reading is prodigious, and even the more familiar of his choices are illuminated by commentaries that crackle with literary and indeed scientific insights", Walter Gratzer, *Nature* **378**, 111 (1995).

Touched with Fire: Manic Depressive Illness and the Artistic Temperament

by Kay Redfield Jamison. Simon and Schuster, £9.99. "An important work that should provoke some serious rethinking in several corners of academe", Hugh Freeman, *Nature* **362**, 666 (1993).

Naturalist by Edward O. Wilson. Penguin, £8.99. "Provides detailed insights into the creation of a renowned scientist, and is an elegant account of the development of his exciting ideas", J. L. Cloudsley-Thomson, *Nature* **372**, 291 (1994).

Lewis Carroll: A Biography by Morton N. Cohen. Papermac, £12. "The author's skill and tact in covering all aspects of Dodgson's complex, enigmatic personality is impressive", Martin Gardner, *Nature* **379**, 127 (1996).

Darwin's Dangerous Idea: Evolution and the Meanings of Life

by Daniel C. Dennett. Penguin, £9.99. "Never in the field of scientific endeavour can so great a theory have been so widely misunderstood by so many with so little reason: but Dennett's book is a marvellous corrective", Mark Ridley, *Nature* **375**, 457 (1995).

Hydrolysis of GTP by elongation factor G drives tRNA movement on the ribosome

Marina V. Rodnina, Andreas Savelsbergh, Vladimir I. Katunin*
& Wolfgang Wintermeyer

Institute of Molecular Biology, University of Witten/Herdecke, D-58448 Witten, Germany

* Sankt Petersburg Nuclear Physics Institute, Russian Academy of Sciences, 188350 Gatchina, Russia

Elongation factor G (EF-G) is a GTPase that is involved in the translocation of bacterial ribosomes along messenger RNA during protein biosynthesis. In contrast to current models, EF-G-dependent GTP hydrolysis is shown to precede, and greatly accelerate, the rearrangement of the ribosome that leads to translocation. Domain IV of the EF-G structure is crucial for both rapid translocation and subsequent release of the factor from the ribosome. By coupling the free energy of GTP hydrolysis to translocation, EF-G serves as a motor protein to drive the directional movement of transfer and messenger RNAs on the ribosome.

THE translocation step of protein elongation entails a large-scale rearrangement of the tRNA-mRNA-ribosome complex¹. Before translocation, deacylated tRNA is bound to the P site (in the P/E hybrid state²) of the ribosome, and peptidyl-tRNA to the A site (A/P state). During translocation, the two tRNAs, together with the mRNA, move synchronously from their pre-translocation to their immediate post-translocation positions, E and P sites, respectively³; from the E site, the deacylated tRNA spontaneously dissociates into solution⁴⁻⁶. In the final post-translocation state, peptidyl-tRNA occupies the P site (P/P state), while E and A sites are unoccupied. Translocation is promoted by elongation factor G (EF-G) which, during the reaction, hydrolyses GTP.

The generally accepted model assumes that EF-G·GTP binds to the pre-translocation ribosome and induces a conformational change which allows translocation, and that, subsequently, EF-G hydrolyses GTP, switches to the low-affinity GDP-bound conformation, and dissociates from the ribosome. The model is based on the observation that EF-G with non-hydrolysable GTP analogues still enhances translocation relative to the spontaneous reaction, but is restricted to a single round⁷⁻⁹ unless the factor is actively removed from the ribosome after translocation⁸. Alternative models in which GTP hydrolysis drives translocation were discussed^{1,10}, but are not generally favoured.

Comparison of the similar crystal structures of EF-G·GDP and nucleotide-free EF-G from *Thermus thermophilus*¹¹⁻¹³ with EF-Tu^{14,15} and with the EF-Tu-aminoacyl-tRNA complex¹⁶ reveals remarkable structural similarities, in particular that between domain IV of EF-G and the anticodon domain of the ternary

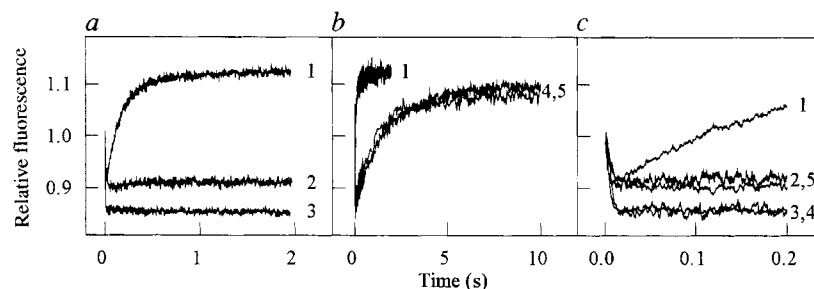
complex. This finding has spurred renewed discussion of the functional parallels and differences of the two factors^{10,15-19}, because an important function of domain IV is suggested by kanamycin-resistance mutations in this domain²⁰, and by the observation that ADP-ribosylation in the homologous region of eukaryotic EF-2 abolishes the factor's function^{21,22}.

Here we present results from both fast kinetic and biochemical experiments in the *Escherichia coli* system which show that EF-G hydrolyses GTP immediately after binding to the ribosome, before, and at least five times faster than, translocation of the tRNAs. When GTP is replaced with a non-hydrolysable analogue, the translocation rate drops more than 50-fold. A similar effect results from the deletion of domain IV. The truncated factor retains full GTPase activity which, however, is restricted to a single round. These findings suggest (1) that GTP hydrolysis by EF-G drives translocation, (2) that domain IV of EF-G has an important role in coupling GTP hydrolysis to translocation, and (3) that domain IV is essential for the release of EF-G·GDP from the ribosome after translocation.

Kinetics of translocation

In vivo, the elongation cycle of *E. coli* proceeds at a rate of 10–20 s⁻¹. Hence fast kinetic techniques have to be applied to resolve individual steps of the cycle and study their mechanisms. Translocation on the ribosome can be monitored by fluorescence stopped-flow using fluorescent tRNA derivatives²³. Here we used a system consisting of highly active components²⁴, in which the partial reactions of elongation proceed at rates consistent with the

FIG. 1 Time course of EF-G-dependent translocation measured by stopped-flow. *a*, Influence of antibiotics. Pre-translocation complex containing deacylated tRNA^{Met} in the P site and fluorescent fMet-Phe-tRNA^{Phe}(Prf16/17) in the A site (0.2 μM) was mixed with EF-G·GTP (1.2 μM) in the absence of antibiotic (trace 1), in the presence of viomycin (trace 2), or of thiostrepton (trace 3). *b*, Translocation with GTP analogues. GTP-free pre-translocation complex (0.2 μM) was mixed with EF-G (4 μM) containing either caged GTP (trace 4) or GDP (trace 5). *c*, EF-G binding in the presence of various guanine nucleotides and antibiotics. Concentrations as for *a*, designation of traces as for *a* and *b*.



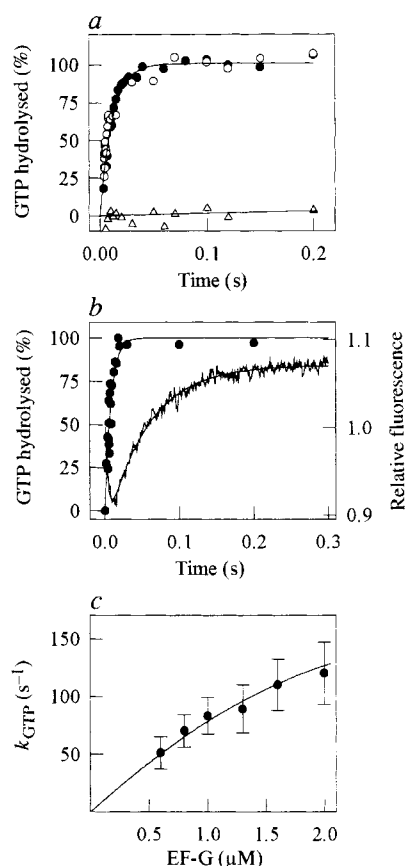


FIG. 2 Kinetics of GTP hydrolysis by EF-G on the ribosome. *a*, Effect of viomycin. Pre-translocation complex (0.1 μM after mixing) was mixed with EF-G[^{32}P]GTP (0.8 μM) without (●) or with (○) viomycin; the control (Δ) was without EF-G. *b*, GTP hydrolysis and translocation at near-saturating concentration of EF-G (2 μM). The fit of the GTPase (●) yielded a rate of 120 s^{-1} , the fluorescence increase was fitted (smooth curve) with $k_{\text{app}} = 25\text{ s}^{-1}$. *c*, Concentration dependence. Rate constants of binding ($k_1 = 140\text{ }\mu\text{M}^{-1}\text{ s}^{-1}$) and dissociation ($k_{-1} = 70\text{ s}^{-1}$) of EF-G GTP and of GTP hydrolysis ($k_{\text{GTP}} = 170\text{ s}^{-1}$) were estimated by fitting (smooth curve) a two-step model⁴⁸.

elongation rate *in vivo*^{25–27}. Pre-translocation ribosomes were formed on near-natural mRNA^{24,28}, such that the A site was occupied with fluorescent fMet-Phe-tRNA^{Phe}(Prf16/17) (proflavin replacing dihydrouracil at positions 16 or 17) and the P site was filled with deacylated initiator tRNA^{fMet}; 80–85% of the complexes were active in EF-G-dependent translocation.

When the pre-translocation complex is rapidly mixed with EF-G·GTP, the fluorescence of fMet-Phe-tRNA^{Phe}(Prf16/17), after a rapid initial drop, increases by about 20% (Fig. 1*a*). The effect is related to translocation, because it is suppressed by adding either viomycin, which inhibits translocation but not GTP hydrolysis²⁹ (see below), or thiostrepton, known to inhibit both translocation and GTP hydrolysis³⁰; indeed, translocation measured by puromycin was completely inhibited by both antibiotics within the time range of the kinetic measurement. The transition is slowed down more than 50-fold (from 25 s^{-1} to 0.5 s^{-1}) when GTP is replaced with the non-hydrolysable GTP analogue, caged GTP (P3-1-(2-nitro) phenylethylguanosine 5'-O-triphosphate³¹), or with GDP (Fig. 1*b*); GTPNP and GTP- γS have a similar effect (not shown).

The rate of the rapid fluorescence decrease is not affected either by replacing GTP with GTP analogues or GDP, nor by adding viomycin or thiostrepton (Fig. 1*c*), although the amplitude is less with GDP and viomycin. Hence we attribute the decrease of the fluorescence signal to the binding of EF-G to the pre-translocation complex. Upon varying the concentration of EF-G

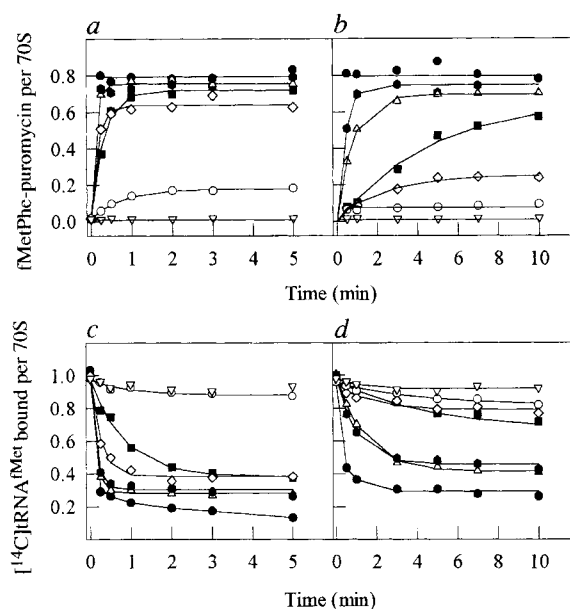


FIG. 3 GTP requirement of translocation. *a*, Puromycin reactivity of fMet-Phe-tRNA^{Phe}, single round; *b*, puromycin reactivity, multiple turnover. *c*, Release of tRNA^{fMet}, single round; *d*, release, multiple turnover. In *a* and *b*, the initiation complex was formed with f[^3H]Met-tRNA^{fMet} and the ternary complex contained *E. coli* [^{14}C]Phe-tRNA^{Phe}; in *c* and *d*, fMet-[^{14}C]tRNA^{fMet} was used for the initiation complex and [^3H]Phe-tRNA^{Phe} for the ternary complex. Concentrations were as follows: pre-translocation complex, 0.1 μM ; EF-G, 0.2 μM (single round) or 0.02 μM (multiple turnover); guanine nucleotides, 1 mM. (●) GTP, (●) caged GTP, (Δ) GTP- γS , (■) GDP, (◇) GTPNP, (○) no nucleotide, (▽) no EF-G.

(up to 2 μM), linear relationships are obtained with GTP, caged GTP, or GDP (data not shown), yielding $k_1 = 120\text{--}150\text{ }\mu\text{M}^{-1}\text{ s}^{-1}$ and $k_{-1} = 50\text{--}150\text{ s}^{-1}$; the same numbers are obtained from the GTPase kinetics (compare with Fig. 2*c*). Apparently, the nucleotide does not appreciably influence the binding the EF-G to the pre-translocation complex.

EF-G-catalysed GTP hydrolysis on the pre-translocation ribosome is very rapid and is not affected when translocation is inhibited by addition of viomycin (Fig. 2*a*). This result can be explained in one of two ways: either GTP hydrolysis takes place before translocation, or the two reactions are not related. The latter alternative is excluded, because GTP hydrolysis accelerates translocation (Figs 1*b* and 3). That EF-G hydrolyses GTP on the ribosome before and independently of translocation is also indicated by the observation that the rate of GTP hydrolysis is the same irrespective of whether the ribosomes are vacant or occupied with tRNA (Table 1). In particular, the fact that GTP hydrolysis is

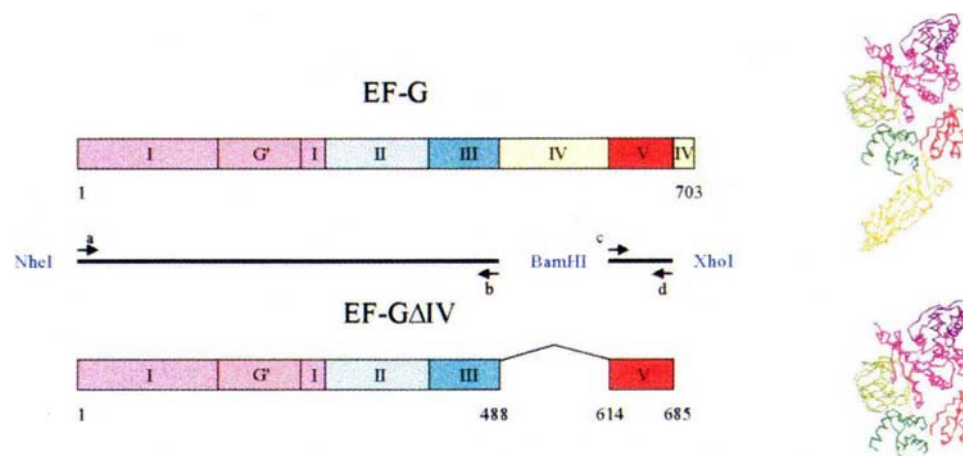
TABLE 1 Rate of GTP hydrolysis by EF-G on tRNA-ribosome complexes

Occupancy of ribosomes	GTPase rate (s^{-1})
Unoccupied	74 ± 23
tRNA ^{fMet} (P)*	109 ± 16
fMet-tRNA ^{fMet} (P)	89 ± 18
tRNA ^{fMet} (P)* Phe-tRNA ^{Phe} (A)	110 ± 24
tRNA ^{fMet} (P), fMet-Phe-tRNA ^{Phe} (A)	71 ± 15

[^{32}P]GTP hydrolysis was measured by quench-flow, mixing equal volumes of ribosomes (0.1 μM after mixing) programmed with MFT-mRNA and of EF-G[^{32}P]GTP (0.8 μM). Ribosomes were either empty or occupied with tRNA^{fMet} or fMet-tRNA^{fMet} in the P site (P) and Phe-tRNA^{Phe} or fMet-Phe-tRNA^{Phe} in the A site (A). The effect of turnover was suppressed by adding unlabelled GTP (1 mM) to the ribosome solution before mixing with EF-G[^{32}P]GTP. Controls without EF-G gave negligible rates.

* Bound without IF2-GTP.

FIG. 4 Domain structures of EF-G and EF-GΔIV. The deletions introduced into *E. coli* EF-G were designed on the basis of the structure of *T. thermophilus* EF-G^{11,12} with appropriate alignments⁴⁹. Domains III and V were connected at amino acids 488 and 614 through a linker of six amino acids (not drawn); the C-terminal amino acids 686 to 703 were also deleted. The structures of EF-G and EF-GΔIV were drawn with RASMOL using coordinates¹¹ of *T. thermophilus* EF-G GDP; domain III is incompletely defined in the structure.



also fast on ribosomes carrying Phe-tRNA^{Phe} in the A site, which is translocated very slowly²³, shows that the GTPase-associated region on the large ribosomal subunit^{30,32–35} is accessible to EF-G independently of occupancy with tRNAs, and supports our conclusion that GTP is hydrolysed before translocation. The reversed sequence of reactions, with translocation preceding³⁵, or limiting³⁶, GTP hydrolysis, is incompatible with our results.

The complete mechanism of translocation comprises four or more steps and cannot yet be described quantitatively. However, the comparison of the rates of GTP hydrolysis and translocation at near-saturating concentration of EF-G (2 μ M; Fig. 2b) reveals that GTP hydrolysis is at least five times faster than translocation (120 s⁻¹ as compared to 25 s⁻¹). From the concentration dependence, a maximum GTPase rate of about 170 s⁻¹ is estimated (Fig. 2c) whereas translocation remains at 25 s⁻¹ (not shown).

GTP hydrolysis for rapid translocation

Guanine nucleotides were also compared in biochemical translocation assays, monitoring both the appearance of puromycin-reactive fMet-Phe-tRNA^{Phe} (Fig. 3a, b) and the release of tRNA^{fMet} from the P site (Fig. 3c, d). The former assay monitors the placement in the P site of the 3' end of fMet-Phe-tRNA^{Phe} on the large ribosomal subunit; the latter reflects the movement of the anticodon domain(s) on the small subunit²⁸.

In the presence of GTP, translocation is unresolvably fast in the

two assays (first measured point at 10 s), both in single turnover (Fig. 3a, c) and in multiple turnover (Fig. 3b, d). With GTP- γ S and caged GTP, the single-round reaction is slower in both assays, although it is still hardly resolved, whereas the turnover reaction with both analogues is appreciably slower. GTPNP is less efficient and apparently limits both assays to a single round. The lack of turnover inhibition observed with two other analogues, caged GTP and GTP- γ S, is difficult to explain; it is not due to hydrolysis as caged GTP is not hydrolysed (according to high-performance liquid chromatography) and neither probably is GTP- γ S. It seems that different analogues differ in functional details. Significant activity of EF-G is also observed in the presence of GDP which, at least in the puromycin assay, eventually leads to complete translocation, albeit at a lower rate, compared with the GTP analogues. By contrast, the release of tRNA^{fMet} is substantially slower with GDP in one round and very slow in turnover. EF-G without nucleotide has very low activity in a single round and practically no activity in turnover.

Deletion of domain IV of EF-G

The activity of truncated EF-G lacking domain IV (EF-GΔIV; Fig. 4) in promoting translocation was monitored by the appearance of puromycin-reactive fMet-[³H]Phe-tRNA^{Phe} and by the release of [¹⁴C]tRNA^{fMet} (Fig. 5a), when EF-GΔIV was found to induce slow translocation on both subunits. The finding that the

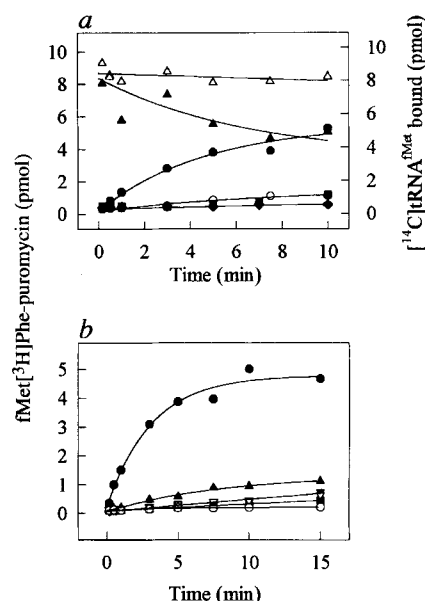


FIG. 5 Activity of EF-GΔIV in translocation. a, Effect of viomycin and thiostrepton. Pre-translocation ribosomes (0.2 μ M) carrying [¹⁴C]tRNA^{fMet} in the P site and fMet-[³H]Phe-tRNA^{Phe} in the A site were mixed with EF-GΔIV (1 μ M). Translocation was measured by both the decrease of ribosome-bound [¹⁴C]tRNA^{fMet} (▲, △) and the increase of puromycin-reactive fMet-[³H]Phe without antibiotic (●, ○), with viomycin (◆), or with thiostrepton (■); controls were without EF-GΔIV (open symbols). b, GTP requirement. Pre-translocation complexes were formed from initiation complex and ternary complex purified from unbound GTP. EF-GΔIV (1 μ M) preincubated with various guanine nucleotides (1 mM) was mixed with an equal volume of pre-translocation complex (0.2 μ M). (●) GTP, (▲) GTP- γ S, (▼) caged GTP, (■) no nucleotide, (○) no EF-GΔIV.

activity is suppressed by adding viomycin or thiostrepton (Fig. 5a), as with EF-G, suggests that the activity of the fragment is related to that of intact factor.

To assess the effect of GTP hydrolysis on the activity of EF-GΔIV in translocation, several non-hydrolysable GTP analogues and GDP were compared with GTP (Fig. 5b). The activity is very low with all nucleotides tested, except for GTP. The difference is probably not due to an affinity effect, because ribosome binding of intact EF-G is about the same with all nucleotides tested (compare with Fig. 1c). Hence the low activity of EF-GΔIV with GTP analogues supports the idea that GTP hydrolysis is required for efficient translocation.

The GTPase activity of EF-GΔIV on the ribosome was tested in both single and multiple turnover (Fig. 6). Under single-turnover conditions (excess factor over pre-translocation complex), EF-GΔIV hydrolyses GTP as rapidly (80 s^{-1}) as EF-G (Fig. 6a). Unlike intact EF-G, however, the truncated factor does not perform multiple turnovers on the ribosome (Fig. 6b), indicating either that it does not dissociate from the ribosome after hydrolysing GTP, or that it does not exchange GDP for GTP in solution. The latter alternative is excluded as pyruvate kinase converts [^3H]GDP to GTP at the same rate, regardless of whether it is bound to EF-G or EF-GΔIV (not shown). Hence the absence of turnover shows that EF-GΔIV does not dissociate from the ribosome after hydrolysing GTP, at least not in the time range of our experiment.

Discussion

According to our results, GTP hydrolysis accelerates translocation considerably relative to the reaction with non-hydrolysable GTP analogues or GDP; the latter is still very significant relative to the factor-free reaction³⁷, which is negligible in our system. There therefore seem to be two ways in which EF-G catalyses translocation. In one, operating with non-hydrolysable GTP analogues or GDP, the activation energy of translocation is reduced by the binding of EF-G to the pre-translocation complex, which is not affected by the nucleotide (Fig. 1c). Significant binding of EF-G GDP to the ribosome has been reported³⁸. Subsequent translocation is relatively slow and the turnover of EF-G even slower,

presumably because the reaction is not driven by GTP hydrolysis and the transition state is reached by thermal motion only. By contrast, in the mechanism operating in the presence of GTP, the hydrolysis of GTP results in a substantial acceleration.

Our data are in partial agreement with published results⁷⁻⁹ in that they confirm that GTP analogues promote translocation and that turnover is inhibited with GTPNP. However, they also show, unlike previous reports but in accordance with the kinetic data, that GTP hydrolysis substantially accelerates translocation to become unresolvably fast both in single and in multiple turnover. The early models of EF-G function in translocation catalysis were based on results obtained with the non-hydrolysable GTP analogues GDPCP and GTPNP, and poly(U)-programmed ribosomes carrying *N*-acetyl-Phe-tRNA^{Phe} in the A site. The assay was done with puromycin, which takes several minutes to react with *N*-acetyl-Phe-tRNA^{Phe} translocated to the P site, precluding kinetic measurements. The results (single-round translocation, no turnover) were interpreted correctly to indicate that GTP hydrolysis is not required for translocation, but for release of EF-G after translocation. The conclusion that GTP hydrolysis must therefore take place after translocation was not justified, however, as no kinetic experiments were done, so the effect of GTP hydrolysis on the rate of translocation could not be observed.

The effect of GTP hydrolysis on the structure of EF-G is unknown, because only the structures of EF-G and EF-G·GDP have been determined^{11,12}. By analogy to the well-characterized structural change of EF-Tu brought about by GTP hydrolysis^{14,15}, the loss of γ -phosphate from GTP probably introduces a conformational strain in domain I of EF-G, which influences the overall architecture of the EF-G molecule; this, in turn, drives the rearrangement of the ribosome. According to our results, domain IV has a role in coupling the conformational change of the factor induced by GTP hydrolysis to the rearrangement of the ribosome leading to translocation. An active movement of domain IV may be involved, perhaps induced by a change in the interaction of domain I with domain V, which is probably rigidly connected to domain IV (ref. 11). The interaction of domains I and V (and II) seems to be important, because at the domain interfaces many fusidic-acid-resistance mutations and revertants have been found^{39,40}. Also, deletion of domain I prevents EF-G from promoting translocation of the tRNAs on the small ribosomal subunit²⁸.

As suggested by the structural homology of domain IV of EF-G with the anticodon arm of the ternary complex ('molecular mimicry'^{16,19}), the main target of domain IV function is likely to be the small ribosomal subunit. It may promote translocation by inducing a rearrangement either within the small subunit or in the relative arrangement of the small and large subunits⁴¹ which leads to a ribosomal structural state ('translocation state') in which the tRNAs can move together with the mRNA. Another possibility is that domain IV, by the movement described, is actively displacing the peptidyl-tRNA from the small ribosomal subunit¹⁰. During or after movement of the peptidyl-tRNA out of the A site, domain IV may move into the A site; this may serve to close the A site for peptidyl-tRNA moving backwards from the translocation transition state.

Unlike EF-G, EF-GΔIV does not dissociate from the ribosome after GTP hydrolysis and translocation, contrary to what might be expected from deleting a domain that probably interacts with the ribosome. Thus domain IV is required not only for rapid translocation but also for release of EF-G. This striking result can be explained by assuming that, early in the sequence of conformational changes induced by GTP hydrolysis, the EF-G-ribosome complex is tightened, enabling the factor to force the ribosome into the transition state of translocation. Movement of domain IV into the A site as described might resolve the tight interaction: for example, a passive movement could induce a structural change in the EF-G molecule that lowers its affinity for the ribosome and aids dissociation; alternatively, an active movement of domain IV might change the conformation of the ribosome to a low-affinity

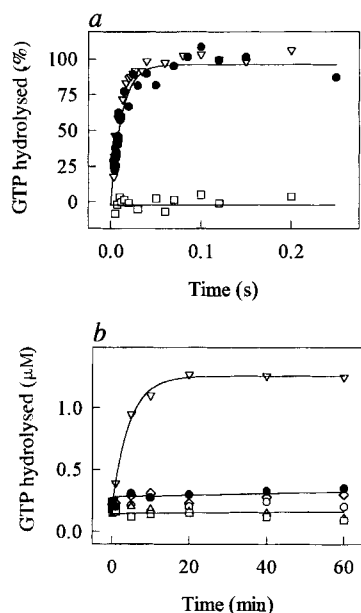


FIG. 6 GTP hydrolysis by EF-GΔIV and EF-G on the ribosome. a, Single round. Pre-translocation complex, 0.1 μM; EF-G and EF-GΔIV, 0.8 μM. b, Multiple turnover. EF-G and EF-GΔIV, 0.03 μM. (▽) EF-G, (●) EF-GΔIV, (◇) ribosomes alone, (△) EF-GΔIV without ribosomes, (○) EF-G without ribosomes, (□) buffer alone.

form for EF-G. EF-GΔIV could not participate in either change and would therefore retain the high affinity gained by GTP hydrolysis and prevent turnover.

Attempts to characterize structural changes of the ribosome during translocation were restricted to a comparison of pre- and post-translocation ribosomes. Scattering methods revealed slight differences^{42,43}, as did visual comparison by cryoelectron microscopy of the pre- and post-translocation states of the ribosome⁴⁴. This lack of any striking difference may be because the structure of the translocating ribosome does not prevail in the post-translocation state, so it is the transition state of translocation that should be studied.

The function of EF-G is to translate chemical energy, derived from GTP hydrolysis, into directional molecular movement on the ribosome, that is, to function as a motor protein. There are similarities in structure between a kinesin-related motor domain and several GTPases, as well as in the switch mechanisms induced by nucleoside triphosphate hydrolysis⁴⁵. The list of proteins related to the classic motor proteins is growing: all promote directional molecular movement at the expense of the free energy of nucleoside triphosphate hydrolysis, including RNA polymerase⁴⁶ and some DNA helicases⁴⁷; EF-G is the first candidate for this list from the GTPase family. □

Methods

Translocation. Experiments were done in buffer A (50 mM Tris-HCl, pH 7.5, 70 mM NH₄Cl, 30 mM KCl, 7 mM MgCl₂, 1 mM dithiothreitol) at 37 °C. Components were prepared as described^{24,28}; viomycin (0.2 mM) and thiostrepton (0.1 mM; 2% DMSO) were from Sigma. Initiation complexes were formed by incubating ribosomes (1 μM) with MFT-mRNA (120 nucleotides; coding sequence AUGUUUACG) (3 μM), *E. coli* f³H]Met-tRNA^{Met} (1.5 μM; 2,250 d.p.m. per pmol), initiation factors 1, 2 and 3 (1.5 μM each), and GTP (1 mM) for 60 min at 37 °C; when necessary, the initiation complex (0.6 ml) was purified

from GTP by gel filtration on Sephacryl S300 (1.6 × 15 cm²). The ternary complex of EF-TuGTP with yeast [¹⁴C]Phe-tRNA^{Phe}(Prf16/17) (1,070 d.p.m. per pmol) was also purified²⁵. Pre-translocation complexes were formed by mixing initiation complex and ternary complex. EF-G-GTP complexes were formed by incubating EF-G (1–4 μM) with GTP (1 mM), phosphoenolpyruvate (3 mM), and pyruvate kinase (5 mg l⁻¹) for 15 min at 37 °C; for experiments with GTP analogues or GDP, nucleotide-free EF-G was incubated with the respective nucleotide (1 mM). EF-G[³²P]GTP was formed by incubating EF-G (1.2–4 μM) with [³²P]GTP (5 μM; 1,000 d.p.m. per pmol), phosphoenolpyruvate (1 mM), and pyruvate kinase (0.5 mg l⁻¹) for 15 min at 37 °C. Complexes of EF-GΔIV were prepared in the same way.

Kinetic experiments. Stopped-flow measurements were done as described²⁵ by rapidly mixing pre-translocation complex and the respective EF-G complex. Assays²⁴ were done in parallel and showed that all [¹⁴C]Phe-tRNA^{Phe}(Prf16/17) added in the ternary complex was bound to the initiation complex (nitrocellulose filter-binding assay) and formed f³H]Met[¹⁴C]Phe-tRNA^{Phe} (peptide analysis by HPLC); in the puromycin assay (1 mM puromycin; 10 s, 37 °C) 80–85% of the ribosome-bound peptidyl-tRNA was translocated. For measuring GTPase rates, equal volumes (26 μl) of the pre-translocation complex (0.2 μM) containing unlabelled GTP (1 mM), and of EF-G or EF-GΔIV (1.2 μM unless indicated) together with [³²P]GTP (5 μM) were rapidly mixed in a quench-flow apparatus (KinTek). [³²P]phosphate was determined as described²⁷. Kinetic data were evaluated by fitting equations containing one or two exponential terms²⁵.

Preparation of EF-GΔIV. PCR mutagenesis was performed with Pfu DNA polymerase (Stratagene) on plasmid pTZ18R containing a DNA fragment encompassing the *E. coli* EF-G gene (*fus*) in the *KpnI* site (pTZfus²⁸). The following primers were used (Fig. 4): a, GAG GAA AGC TAG CGC TCG TAC AA; b, TTT CGG ATC CAG CAA CCT GCG GTT T; c, AAC CAG TTC TGG GAT CCC CGA TCA T; d, GCG CCT CGA GTT ACT TCA GGA AT. PCR products were cleaved with *NheI*/*Bam*HI and *Bam*HI/*XhoI*, respectively, ligated through the *Bam*HI site and cloned into the *NheI* and *XhoI* sites of pRSET C (Invitrogen). The construct was verified by DNA sequencing using the Sequenase 2.0 kit (US Biochemicals). EF-GΔIV, which at the N terminus contained an additional 11 amino acids including six histidines, was expressed in *E. coli* BL21(DE3)pLysS (Novagen) and purified from inclusion bodies by affinity chromatography on Ni²⁺-NTA-agarose in urea (Quiagen). Concentration and rebuffing into buffer A was by ultrafiltration (Amicon). Purified EF-GΔIV did not contain detectable amounts of intact EF-G, as shown by immunoblot analysis²⁸ of an overloaded SDS gel (not shown).

Received 4 October; accepted 18 November 1996.

- Spirin, A. S. *Prog. Nucl. Acids Res. Mol. Biol.* **32**, 75–114 (1985).
- Moazed, D. & Noller, H. F. *Nature* **342**, 142–148 (1989).
- Paulsen, H. & Wintermeyer, W. *Biochemistry* **25**, 2749–2756 (1986).
- Spirin, A. S. *FEBS Lett.* **165**, 280–284 (1984).
- Robertson, J. M. & Wintermeyer, W. *J. Mol. Biol.* **196**, 525–540 (1987).
- Semenkov, Y. P., Rodnina, M. V. & Wintermeyer, W. *Proc. Natl Acad. Sci. USA* **93**, 12183–12188 (1996).
- Inoue-Yokosawa, N., Ishikawa, C. & Kaziro, Y. *J. Biol. Chem.* **249**, 4321–4323 (1974).
- Belitsina, N. V., Glukhova, M. A. & Spirin, A. S. *FEBS Lett.* **54**, 35–38 (1975).
- Modollel, J., Girbes, T. & Vazquez, D. *FEBS Lett.* **60**, 109–113 (1975).
- Abel, K. & Jurmak, F. *Structure* **4**, 229–238 (1996).
- Czworkowski, J., Wang, J., Steitz, T. A. & Moore, P. B. *EMBO J.* **13**, 3661–3668 (1994).
- Aevansson, A. et al. *EMBO J.* **13**, 3669–3677 (1994).
- Al-Karadaghi, S., Aevansson, A., Garber, M., Zheltonosova, J. & Liljas, A. *Structure* **4**, 555–564 (1996).
- Berchtold, H. et al. *Nature* **365**, 126–132 (1993).
- Kjeldgaard, M., Nissen, P., Thirup, S. & Nyborg, J. *Structure* **1**, 35–50 (1993).
- Nissen, P. et al. *Science* **270**, 1464–1472 (1995).
- Liljas, A. et al. *Biochem. Cell Biol.* **73**, 1209–1216 (1995).
- Moore, P. B. *Science* **270**, 1453–1454 (1995).
- Nyborg, J. et al. *Trends Biochem. Sci.* **21**, 81–82 (1996).
- Hou, Y., Lin, Y. P., Sharer, J. D. & March, P. E. *J. Bacteriol.* **176**, 123–129 (1994).
- Nygård, O. & Nilsson, L. *Biochim. Biophys. Acta* **824**, 152–162 (1985).
- Omura, F., Kohno, K. & Uchida, T. *Eur. J. Biochem.* **180**, 1–8 (1989).
- Robertson, J. M., Paulsen, H. & Wintermeyer, W. *J. Mol. Biol.* **192**, 351–360 (1986).
- Rodnina, M. V. & Wintermeyer, W. *Proc. Natl Acad. Sci. USA* **92**, 1945–1949 (1995).
- Rodnina, M. V., Fricke, R. & Wintermeyer, W. *Biochemistry* **33**, 12267–12275 (1994).
- Rodnina, M. V., Fricke, R., Kuhn, L. & Wintermeyer, W. *EMBO J.* **14**, 2613–2619 (1995).
- Rodnina, M. V., Pape, T., Fricke, R., Kuhn, L. & Wintermeyer, W. *J. Biol. Chem.* **271**, 646–652 (1996).
- Borowski, C., Rodnina, M. V. & Wintermeyer, W. *Proc. Natl Acad. Sci. USA* **93**, 4202–4206 (1996).
- Modollel, J. & Vazquez, D. *Eur. J. Biochem.* **81**, 491–497 (1977).
- Cundliffe, E. in *The Ribosome. Structure, Function, and Evolution* (eds Hill, W. E. et al.) 479–490 (Am. Soc. Microbiol., Washington DC, 1990).
- Schlichting, I. et al. *Proc. Natl Acad. Sci. USA* **86**, 7687–7690 (1989).
- Moazed, D., Robertson, J. M. & Noller, H. F. *Nature* **334**, 362–364 (1988).
- Ryan, P. C., Lu, M. & Draper, D. E. *J. Mol. Biol.* **221**, 1257–1268 (1991).
- Rosendahl, G. & Douthwaite, S. *Nucleic Acids Res.* **23**, 2396–2403 (1995).
- Kaziro, Y. *Biochim. Biophys. Acta* **505**, 95–127 (1978).
- Robertson, J. M., Urbanke, C., Chinali, G., Wintermeyer, W. & Parmeggiani, A. *J. Mol. Biol.* **189**, 653–662 (1986).
- Gavrilova, L. P., Kostishchikina, O. E., Kotliansky, V. E., Rutkevitch, N. M. & Spirin, A. S. *J. Mol. Biol.* **101**, 537–552 (1976).
- Baca, O. G., Rohrbach, M. S. & Bodley, J. W. *Biochemistry* **15**, 4570–4574 (1976).
- Johanson, U. & Hughes, D. *Gene* **143**, 55–59 (1994).
- Johanson, U., Aevansson, A., Liljas, A. & Hughes, D. *J. Mol. Biol.* **258**, 420–432 (1996).
- Spirin, A. S. *Cold Spring Harb. Symp. Quant. Biol.* **34**, 197–207 (1969).
- Spirin, A. S. *Biochimie* **69**, 949–956 (1987).
- Serdyuk, I. et al. *Biochimie* **74**, 299–306 (1992).
- Stark, H. et al. *Cell* (in the press).
- Sablin, E. P., Kull, F. J., Cooke, R., D. & Fletterick, R. J. *Nature* **380**, 555–559 (1996).
- Yin, H. et al. *Science* **270**, 1653–1657 (1995).
- West, S. C. *Cell* **86**, 177–180 (1996).
- Bernasconi, C. F. *Relaxation Kinetics* (Academic, New York, 1976).
- Aevansson, A. *J. Mol. Evol.* **41**, 1096–1104 (1995).

ACKNOWLEDGEMENTS. We thank A. Spirin and H. Noller for comments on the manuscript; G. Gualerzi for gifts of initiation factors and mRNA constructs; R. Goody for caged GTP; and P. Striebeck for technical assistance. The work was supported by the Deutsche Forschungsgemeinschaft and the Volkswagen-Stiftung.

CORRESPONDENCE and requests for materials should be addressed to W.W. (e-mail: winterme@uni-wh.de).

Dynamical evolution of Jupiter's Trojan asteroids

Harold F. Levison*, Eugene M. Shoemaker†
& Carolyn S. Shoemaker‡

* Space Science Department, Southwest Research Institute, Boulder, Colorado 80302, USA

† Lowell Observatory, Flagstaff, Arizona 86001, USA

‡ Northern Arizona University, Flagstaff, Arizona 86001, USA

TROJAN asteroids, which may outnumber the asteroids in the asteroid belt, are objects that orbit the Sun with the same mean semi-major axis as Jupiter, but lead or trail the position of Jupiter in its orbit by $\sim 60^\circ$. One very interesting aspect of the Trojan swarms is that a significant number of asteroids are on orbits that analytic theory suggests should be unstable¹. Here we present the results of long-term dynamical integrations of the Trojan asteroids, that enable us to investigate the stability of the swarm population. We find that the orbits of the swarm asteroids are not stable indefinitely—the gravitational effects of the giant planets have reduced the swarms' outer boundaries over time. We estimate that there are over 200 escaped Trojan asteroids with diameters >1 km currently roaming the Solar System, a few of which may be on Earth-crossing orbits.

The Trojan asteroids orbit the Sun about the L_4 and L_5 Lagrange points of Jupiter, and are locked in a 1:1 mean motion resonance with Jupiter. Early analytic work on the three-body problem (Sun–Jupiter–asteroid) has shown that asteroids at the L_4 and L_5 points are in stable orbits, and that objects slightly displaced from these points will oscillate about them². The dynamical behaviour of these asteroids has been widely studied by both analytical means (for example, refs 3 and 4) and using relatively short-term numerical integrations^{5–7}.

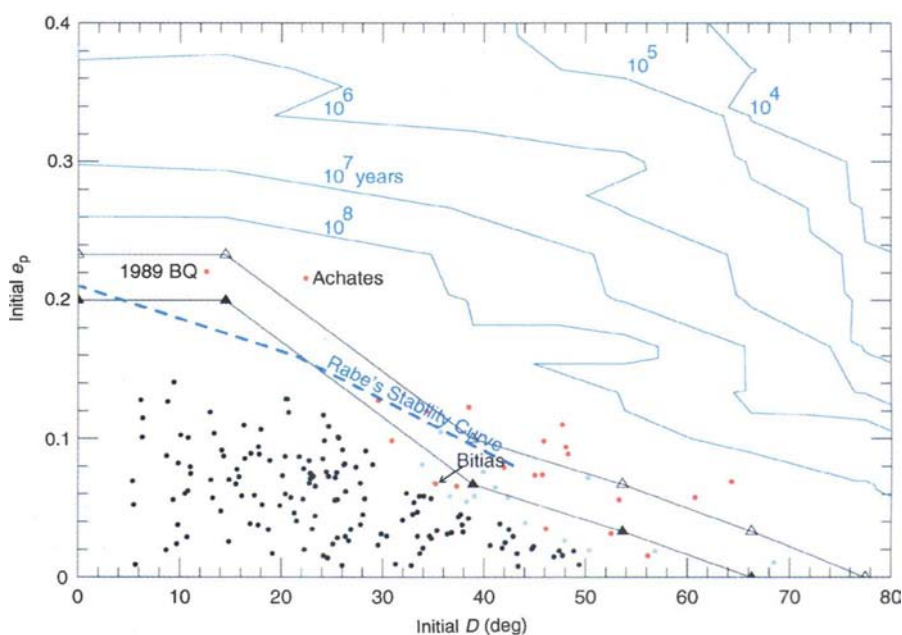
Most Trojan asteroids are probably the remains of a population of planetesimals that formed the core of Jupiter¹. In addition, as they are dynamically dominated by Jupiter, understanding the dynamical structure of the Trojan swarms may lead to important clues to the origin and early dynamical evolution of this planet. Perhaps the most important aspect of the structure of the swarms is their outer boundaries. The only prior attempt to understand the outer boundary of the swarms was by Rabe⁸, using analytical techniques. However, a comparison between his result and the distribution of real Trojans (Fig. 1) shows that a significant number of real Trojans lie above Rabe's stability curve. (It should be noted that these Trojans were discovered after Rabe published this work.)

The discrepancy between the observations and theory prompted Shoemaker, Shoemaker and Wolfe¹ to argue that the true stability curve lies above the curve calculated by Rabe. Their argument was based on the long-held assumption by the astronomical community that the Solar System and structures in it are dynamically stable. However, recent results^{9–11} have shown that chaos, caused by the overlap of resonances within the Solar System, can lead to behaviour of small bodies that appears stable for billions of years after which a global instability may occur. These results call into question both the conclusions of ref. 1 and the nature of the outer boundary of the Jupiter's Trojan swarms.

In an attempt to understand the origin of the objects above Rabe's stability curve, we first numerically integrated the orbits of 270 fictitious, massless L_4 Trojan asteroids in three dimensions under the influence of the Sun and the four giant planets; the particles themselves were not gravitationally interacting with each other. The equations of motion for the Sun, planets and particles were integrated using the very efficient MVS integration scheme¹², with a timestep of 1/2 year. We integrated the orbits of the test particles for 10^9 years or until they escaped the swarms.

Each of our fictitious Trojans had different initial values of libration amplitude, D , and proper eccentricity, e_p (see Fig. 1

FIG. 1 Isoleths of the dynamical lifetimes for Jupiter Trojans. Dots show the location of the known Trojan asteroids as a function of their proper elements. Owing to secular perturbations, the osculating orbital elements vary with time. However, it is possible to develop a set of proper orbital elements that remain approximately constant¹³. Trojan asteroids librate about points that are $\sim \pm 60^\circ$ from Jupiter. (The actual location of these fixed points are a function of the proper eccentricity¹⁸.) If ϕ is defined as the instantaneous difference between the mean longitudes of the asteroid and Jupiter, then the libration amplitude, D , is the amplitude of oscillations in ϕ . The proper eccentricity, e_p , is defined as the amplitude of the oscillation in $e \cos(\tilde{\omega} - \tilde{\omega}_J)$, where e is the eccentricity of the asteroid and $\tilde{\omega}$ and $\tilde{\omega}_J$ are the longitudes of perihelia for the asteroid and Jupiter, respectively. The dashed blue curve shows the location of Rabe's⁸ analytic stability limit. Notice that there are significant differences between Rabe's prediction and the observed location of the upper boundary of the swarms. Of particular interest here are the Trojans that are above Rabe's curve. Shoemaker et al.¹ used these objects (except 1989 BQ and Achates which were not known at that time) to conclude that the true stability curve lies above Rabe's. The solid light blue curves are contours of the dynamical lifetimes of particles in our first set of integrations, which were for 1 Gyr. The stable region shrinks quite dramatically over this time span. It is not possible to plot a contour at 10^9 years because the main integration was stopped at this point. To determine the limits for this isopleth, for each D we plot an open triangle showing the orbit with the smallest e_p that was unstable in 1 Gyr; the filled triangle



represents with the largest e_p that was stable. The 10^9 -year lifetime isopleth must lie between these two extremes. The colour of a dot representing a known Trojan indicates whether the object is stable (green) or unstable (red) as determined by our second set of integrations; objects whose orbits we did not integrate are shown in black.

legend). These initial proper elements were chosen from a grid of 27 values of e_p (ranging from 0 to 0.8) and 10 values of D (between 0° and 140°). In addition, associated with both D and e_p , which are amplitudes of values that oscillate as a function of time, are two phases, which were chosen to be the same for all 270 particles. However, through a preliminary set of calculations, we find that the dynamical lifetimes of particles are not significantly affected by our choice of the phases. The test particles initially started in the same plane as Jupiter, but the planets had their correct inclinations and the calculation was performed in three dimensions.

We were interested in whether an orbit is stable for the length of the integration. An orbit was deemed unstable if the particle either suffered a close approach with Jupiter and was ejected from the Solar System or visited all four quadrants of a coordinate frame that rotates with the Sun–Jupiter system. If a particle visited all four quadrants of such a frame then it was either in a non-resonant or a horseshoe orbit and was no longer a member of the Trojan swarms.

The results from this initial study are shown in Fig. 1 as contours of dynamical lifetime. The integrations show that the observed

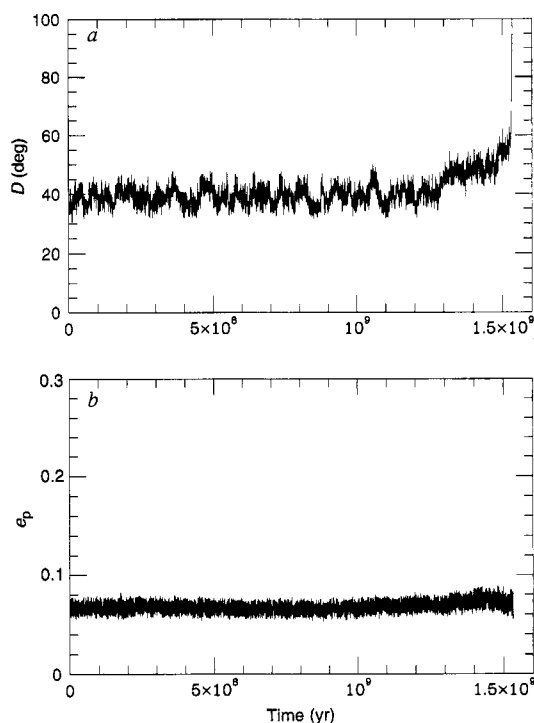


FIG. 2 Temporal variation of D (panel a) and e_p (panel b) of a typical Trojan on an unstable orbit, (5120) Bitias (see Fig. 1). The value of e_p is approximately constant while D increases as a function of time. Note that D remains approximately constant for over a billion years and starts to grow only near the end of the asteroid's life as a Trojan. This behaviour is typical of the evolution of Trojans as all 21 unstable Trojans investigated behaved in a similar way. It is also apparently typical of chaotic orbits in other regions of the Solar System (see ref. 19 for examples in the Kuiper belt). It also explains why most of the unstable Trojans' proper eccentricities are within the same range of e_p found for stable Trojans. We infer that many of the unstable Trojans initially had smaller values of D and evolved onto their current orbits in the process of leaving the swarm. Additionally, many of the unstable Trojans may have been displaced in D from more stable orbits by collisions¹. If indeed e_p remains approximately constant, it is difficult to understand the origin of the unstable objects (4835) 1989BQ and (5144) Achates, which have large e_p (Fig. 1). There are two possible explanations for these high-eccentricity orbits. (1) These objects initially had very small values of D and we are currently observing them as they slowly evolve out of the swarms. This seems unlikely as there are no other known objects in the swarms with such large values of e_p . (2) These high- e_p orbits may be the result of collisions.

upper boundary in e_p – D space occupied by Jupiter Trojans is time dependent. In times past, it was significantly larger than at present and it has been slowly shrinking with time. Indeed, the area of physical space inhabited by Trojan asteroids shrinks by almost a factor of two between 10^8 and 10^9 years. Another result from our billion-year survey is that a significant number of the present Trojans lie above or to the right of our billion-year stability curve. About 10% of the Trojans lie in a region that is unstable on the billion-year timescale while another 10% fall in the region between the lines marked by open and closed triangles in Fig. 1.

To better understand and test the observation that a significant number of Trojans are outside the billion-year stability region, we numerically integrated the orbits of 36 real Trojan asteroids near the stability boundary. These Trojans are indicated in Fig. 1 by either green or red dots. (The current orbital elements for the Trojans and the planets were provided by E. Howell and L. Wasserman.) The orbit of each Trojan was integrated for up to 4×10^9 years, or until it suffered a close encounter with one of the planets or was ejected from the Solar System.

Of the 36 objects studied, 21 were unstable in less than 4 billion years. The objects we found to be unstable are represented by red dots in Fig. 1, while those that were stable are shown in green. (Black dots represent Trojans that we did not integrate.) All the objects with current proper elements above the curve connecting the open triangles (those objects in our initial survey that were unstable after 10^9 years), were indeed unstable. In addition, all particles above Rabe's⁸ analytical stability curve also were unstable. All the unstable Trojans followed similar behaviours, as illustrated in Fig. 2.

Our main finding is that the Trojan swarms of Jupiter are slowly dispersing. Indeed, our results show that the Trojan swarms must be added to the growing list of dynamically unstable structures in the Solar System.

One interesting side issue are the fates of the objects that escape the swarms. To address this, we integrated their orbits under the gravitational influence of the Sun and the four giant planets using the RMVS integration scheme¹³. Our integration scheme is based on the MVS method described in ref. 12, but it can handle close encounters between particles and planets. The trajectory of each escaped Trojan was followed for 2×10^7 years unless it either was ejected from the Solar System, or struck the Sun or a planet.

We found that, as first suggested by Rabe⁸, escaped Trojans evolve onto orbits similar to those of Jupiter-family comets¹⁴ (Fig. 3). It is pertinent to ask how many, if any, of the observed Jupiter-family comets are escaped Trojans, assuming that Trojans can become active like comets. A total of 178 known Trojans are shown in Fig. 1. Of these, 21 (12%) are unstable over a 4-billion-year time interval, assuming that the ones that we did not integrate are stable. This assumption is probably valid as the remaining objects are well within our derived stability limit. The probability, p , that any given object will leave the swarm in a year is $0.12/4 \times 10^9 = 2.9 \times 10^{-11} \text{ yr}^{-1}$. We estimate that there are $\sim 2 \times 10^6$ Jupiter Trojan asteroids with radii greater than 1 km, the typical size of a comet¹. Thus, the leakage rate of bodies with diameters greater than 1 km from the Trojan swarms due to dynamical erosion is $r = 6.2 \times 10^{-5} \text{ yr}^{-1}$. Collisions may also play a role in knocking Trojans out of the swarm. Marzari *et al.*¹⁵ have estimated that r due to collisions is 10^{-3} yr^{-1} , significantly higher than the rate from dynamical erosion.

The average number of escaped Trojans still resident in the Solar System is $N_{\text{tot}} = rL$, where L is the mean time that an escaped object remains in the Solar System and r is the sum of the rates due to collisions and dynamical erosion. From our integrations $L = 1.2 \times 10^6$ years, thus $N_{\text{tot}} = 1,200$. However, these objects are spread throughout the Solar System. Only $\sim 1\%$ have perihelia small enough to be easily discovered (defined as $q < 2.5 \text{ AU}$), which implies that no more than about 12 of the known Jupiter-family comets are escaped Trojans. The expected number that might be recognizable as comets is still less because the physical lifetimes of comets (the time over which they can

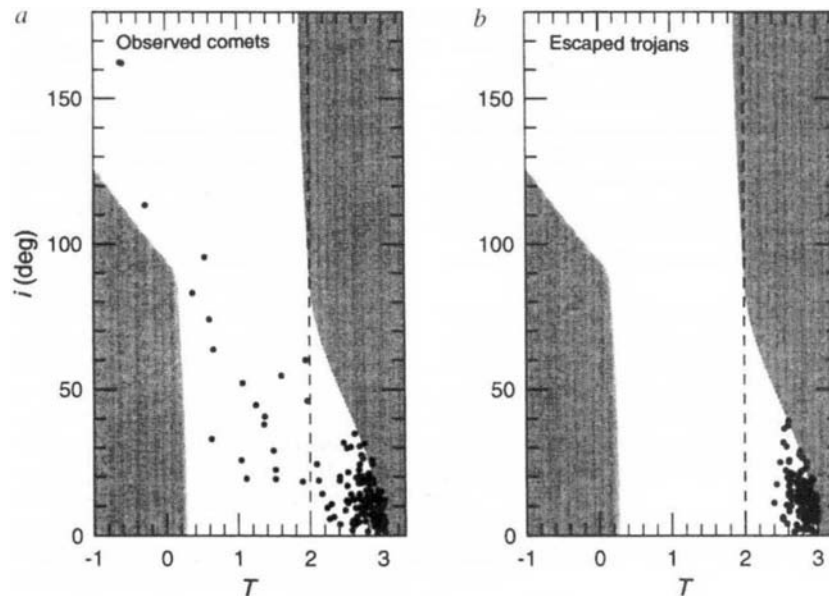


FIG. 3 Comparison between the orbital elements of the observed comets with periods less than 200 years and those calculated for escaped Trojans. *a*, The inclinations (i) of the known periodic comets versus their Tisserand parameters (T). The Tisserand parameter is usually defined as $T = (a_j/a) + 2\sqrt{(1-e^2)(a/a_j)} \cos i$ where a_j is Jupiter's semi-major axis and a , e and i are the comet's semi-major axis, eccentricity and inclination, respectively. It is an approximation to the Jacobi constant, which is an integral of the motion in the circular restricted three-body problem. It is also a measure of the relative velocity between a comet and Jupiter during close encounters, $v_{\text{rel}} \propto \sqrt{3-T}$. Jupiter-family comets are defined¹⁴ as those with $T > 2$; they lie to the right of the dashed line in the figure. This limit on T

implies that Jupiter-family comets have small encounter velocities with Jupiter. The shaded areas represent regions that are physically unattainable for cometary orbits. *b*, The inclination and Tisserand parameter of the escaped Trojans asteroids. Each object is plotted once when it first evolves onto an orbit with a perihelion distance $< 2.5 \text{ AU}$, where it is assumed that the object, if it is active, could be discovered by the usual methods of telescopic detection (see refs 16, 20). The discovery of comets with $q > 2.5 \text{ AU}$ is far from complete. This figure shows that the orbital element distribution of recently escaped Trojans is very similar to that of the known Jupiter-family comets.

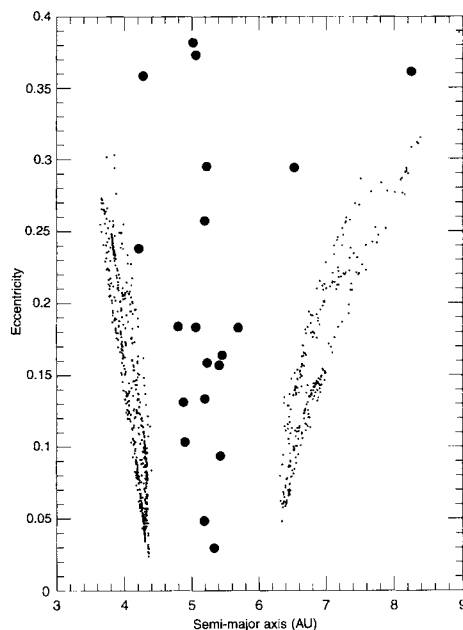


FIG. 4 A comparison between the orbital elements of escaped Trojans just after they first encounter Jupiter (large dots) and the possible pre-capture orbits of the comet Shoemaker-Levy 9 (small dots) as determined by Chodas and Yeomans²¹. The figure shows that there is very little if any overlap between the two distributions, indicating that Shoemaker-Levy 9 was not likely to have been an escaped Trojan.

produce a coma and tail) is much less than their dynamical lifetime, because only about 5–20% of comets are active^{16,17}. If escaped Trojans behave like typical comets, then the total number of active Jupiter-family comets that are escaped Trojans is only ~ 1 out of the 150 known Jupiter-family comets.

Was the comet Shoemaker-Levy 9, which struck Jupiter in 1994, an escaped Trojan asteroid of Jupiter? Figure 4 shows that this is quite unlikely, because the orbital elements of recently escaped Trojans do not match the probable pre-capture orbital elements of Shoemaker-Levy 9. $\square$

Received 18 July; accepted 22 October 1996.

- Shoemaker, E. M., Shoemaker, C. S. & Wolfe, R. F. In *Asteroids II* (eds Benz, R. P., Gehrels, T. & Matthews, M. S.) 921–948 (Univ. Arizona Press, Tucson, 1989).
- Szebehely, V. *Theory of Orbits: The Restricted Problem of Three Bodies* (Academic, London, 1967).
- Erdi, B. *Celest. Mech.* **18**, 141–161 (1978).
- Rabe, E. *Astron. J.* **70**, 687–688 (1965).
- Bien, R. & Schubart, J. *Astron. Astrophys.* **175**, 292–298 (1987).
- Mikkola, S. & Innanen, K. *Astron. J.* **104**, 1641–1649 (1992).
- Milani, A. *Celest. Mech.* **57**, 59–94 (1993).
- Rabe, E. *Astron. J.* **72**, 10–17 (1967).
- Wisdom, J. *Icarus* **56**, 51–74 (1983).
- Torbett, M. *Astron. J.* **98**, 1477–1481 (1989).
- Duncan, M. J., Levison, H. & Budd, M. *Astron. J.* **110**, 3073–3081 (1995).
- Wisdom, J. & Holman, M. *Astron. J.* **102**, 1528–1538 (1991).
- Levison, H. & Duncan, M. *Icarus* **108**, 18–36 (1994).
- Levison, H. in *Completing the Inventory of the Solar System* (ed. Nettig, T. W. & Hahn, J. M.) 173 (ASP Conf. Ser. 107, Astron. Soc. Pac., San Francisco, 1997).
- Marzari, F., Farinella, P., Davis, D. R., Scholl, H. & Campo Bagatin, A. *Icarus* (submitted).
- Levison, H. & Duncan, M. *Icarus* (submitted).
- Shoemaker, E. M., Weissman, P. R. & Shoemaker, C. S. in *Hazards Due to Comets and Asteroids* (ed. Gehrels, T.) 131 (Univ. Arizona Press, Tucson, 1994).
- Levison, H., Shoemaker, E. M. & Wolfe, R. F. *Lunar Planet. Sci. Conf.* **XXII**, 803–804 (1991).
- Levison, H. & Duncan, M. *Astrophys. J.* **406**, L35–L38 (1993).
- Duncan, M., Quinn, T. & Tremaine, S. *Astrophys. J.* **328**, L69–L73 (1988).
- Chodas, P. & Yeomans, D. in *The Collision of Comet Shoemaker-Levy 9 and Jupiter* (eds Noll, K. S., Feldman, P. D. & Weaver, H. A.) (Cambridge Univ. Press, in the press).

ACKNOWLEDGEMENTS. We thank C. Chapman, M. Duncan and A. Stern for discussions. H.F.L. was supported by the NASA Origins of Solar Systems Program and the NASA Planetary Geology and Geophysics Program.

CORRESPONDENCE should be addressed to H.F.L. (e-mail: hal@gort.space.swri.edu).

Ray and wave chaos in asymmetric resonant optical cavities

Jens U. Nöckel & A. Douglas Stone

Department of Applied Physics, Yale University, New Haven, Connecticut 06520-8284, USA

OPTICAL resonators are essential components of lasers and other optical devices. A resonator is characterized by a set of modes, each with a resonant frequency ω and resonance width $\delta\omega = 1/\tau$, where τ is the lifetime of a photon in the mode. Cylindrical or spherical dielectric resonators have extremely long-lived resonances¹ due to 'whispering gallery' modes in which light circulates around the perimeter trapped by total internal reflection. These resonators emit light isotropically. Recently a new category of asymmetric resonant cavities has been proposed in which substantial deformation of the cavity from cylindrical or spherical symmetry leads to partially chaotic ray dynamics. This has been predicted²⁻⁴ to give rise to a universal, frequency-independent broadening of the whispering-gallery resonances, and to highly anisotropic emission. Here we present solutions of the wave equation for asymmetric resonant cavities which confirm these predictions but also reveal interesting frequency-dependent effects characteristic of quantum chaos. For small deformations the lifetime is controlled by evanescent leakage, the optical analogue of quantum tunnelling⁵; here the lifetime is significantly shortened by a process known as 'chaos-assisted tunnelling'^{6,7}. In contrast, for large deformations ($\sim 10\%$) some resonances are found to have longer lifetimes than predicted by the ray chaos model due to the phenomenon of 'dynamical localization'⁸.

The prediction of universal behaviour of the whispering-gallery resonances is derived from the limit of ray optics in which the wavelength of the light, λ , is much shorter than the radius of curvature of the asymmetric resonant cavity (ARC). Here for simplicity we will focus on the effectively two-dimensional case of a deformed cylindrical resonator as in Fig. 1, top; the case of deformed spheres (for example, liquid droplets) in the ray limit

has been discussed elsewhere³. When the cylinder is undeformed (circular), the angle of incidence χ is conserved at each collision and escape occurs isotropically by the exponentially slow process of evanescent leakage. When the cylinder is sufficiently deformed a new decay process, refractive escape, becomes possible^{2,4} in which a ray initially in a whispering-gallery trajectory diffuses chaotically in phase space until it reaches the critical angle $\chi_c = \sin^{-1}(1/n)$, where n is the refractive index of the dielectric, and is refracted out of the resonator. This is illustrated schematically in Fig. 1, bottom. As no refractive escape is allowed for $\chi > \chi_c$, the ray dynamics is equivalent to the nonlinear dynamics of a point mass undergoing specular reflection from the walls of a two-dimensional "billiard"⁹. For smooth convex deformations from a circle, this dynamics becomes increasingly chaotic with increasing deformation according to the Kolmogorov–Arnold–Moser (KAM) scenario^{4,10,11}. One stage in this evolution is shown in the Poincaré surface of section (Fig. 2a). Initially, small regions of chaotic behaviour appear near unstable periodic orbits which

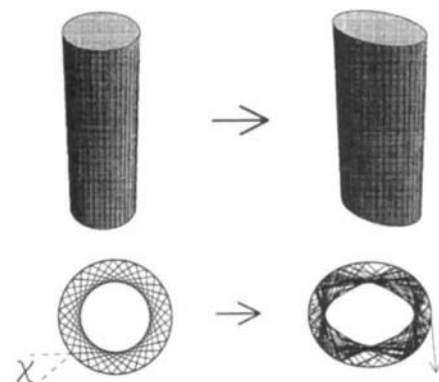


FIG. 1 Top, deformation of a dielectric cylinder of radius R to make an asymmetric resonant cavity (ARC). All the results below are for a quadrupolar deformation of the cross-section, so that the radius as a function of angle $r(\phi)$ is given by $r_0(1 + \varepsilon \cos 2\phi)$, where $\varepsilon < 0.2$ measures the size of the deformation and $r_0(\varepsilon)$ is chosen to maintain a fixed area. Bottom, regular and chaotic ray trajectories in the plane perpendicular to the cylinder axis for a whispering-gallery orbit for $\varepsilon = 0$ (left) and $\varepsilon = 0.1$ (right); the chaotic trajectory eventually escapes refractively when the sine of the angle of incidence, $\sin \chi$, is $< 1/n$.

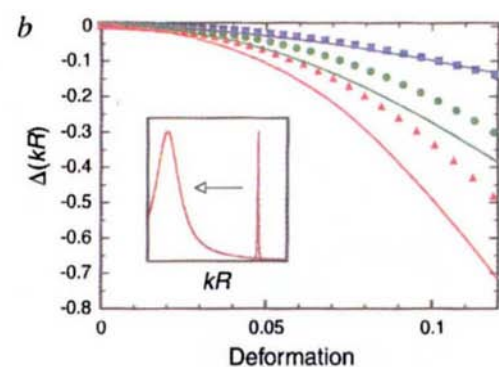
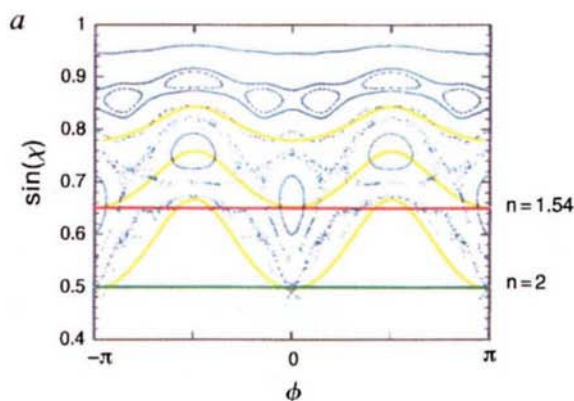


FIG. 2. a, Partial Poincaré surface of section showing phase-space ray dynamics for several trajectories for $\varepsilon = 0.072$. The $\sin \chi$ and the angular position ϕ are recorded at each collision. The green and red horizontal lines represent the critical lines $\sin \chi = 1/n$ for refractive escape for $n = 2, 1.54$. 'Islands' indicate trajectories oscillating around stable periodic orbits. Unbroken KAM curves appear above $\sin \chi \approx 0.8$; no whispering-gallery orbit can cross such a curve. Trajectories below $\sin \chi \approx 0.8$ lying outside islands are all chaotic, but when they are followed for 100–200 reflections (as above) they remain near the adiabatic KAM curves predicted by equation (1) (yellow lines), except when islands intersect the relevant curve (see the discussion of peak-splitting in Fig. 4 legend). b, Shift in

resonance frequency with deformation ε from exact numerical solution of the wave equation for the ARC resonances (solid lines) and from eikonal approximation for the bound states (symbols, adjusted to agree at $\varepsilon = 0$). Units of frequency are $kR = 2\pi R/\lambda$ where λ is the wavelength outside the dielectric. For $\varepsilon = 0$, the resonances are at $kR \approx 12.1, 27.5, 44.6$ (blue, green, red). Inset, schematic showing the generic redshift and broadening of these resonances with deformation. The redshift may be understood as follows: the approximate condition for a whispering-gallery resonance is that an integer number of wavelengths fit around the perimeter (see Fig. 4a). As the ARC is deformed, for a fixed area the perimeter increases requiring the wavelength to increase.

alternate with stable orbits and their associated islands, but any chaotic trajectory is confined to a small separatrix region near the islands by KAM curves (which correspond to families of regular quasi-periodic orbits). As the deformation increases these KAM curves break up, allowing refractive escape of whispering-gallery orbits. Because ray optics describes the limit $\lambda \rightarrow 0$, the escape rate due to this process is independent of λ and should be the same for all resonances corresponding to the same set of whispering-gallery orbits.

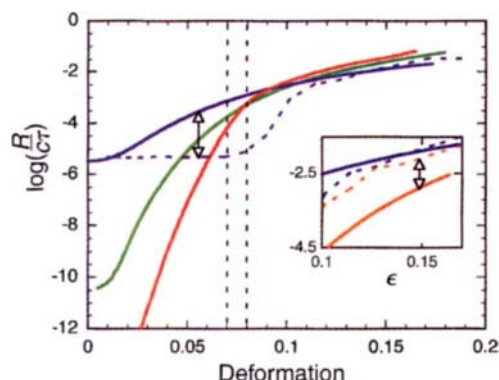


FIG. 3 Logarithm of inverse resonance lifetimes (in units of R/c) versus deformation ϵ for the same resonances as in Fig. 2b (same colour coding), determined numerically. All three resonances correspond to an initial adiabatic curve with initial $\sin \chi \approx 0.8$. The dashed blue line is the result for this adiabatic curve from the ray-optics model for the $kR = 12.1$ resonance. The escape probability for collisions at $\sin \chi$ is determined by the semiclassical relation $\sin \chi = m/nkR$ and the known resonance lifetimes of the cylinder for angular momentum m (J.U.N. *et al.*, manuscript in preparation). Dashed black lines delimit the cross-over region from evanescent to refractive escape when the last intervening KAM curve breaks; as expected all three resonances have approximately the same lifetime above this deformation. The discrepancy at small ϵ (vertical line with arrowheads) is indicative of chaos-assisted tunnelling. Inset, expanded scale showing the resonance with $kR \approx 12.1$ and its agreement with the ray-optics approximation (dashed blue) at high deformations. Also shown (brown) is a resonance with initial $\sin \chi \approx 0.9$, $kR \approx 33.2$, which has a substantially longer lifetime than predicted by the ray-optics model (dashed brown). The difference marked by the line with arrowheads indicates the dynamical localization of photons.

gallery orbits.

To compare the previous model derived from ray optics^{2,4} to solutions of the wave equation we have to overcome a fundamental problem. In regular (integrable) systems a set of trajectories on a KAM curve corresponds to a set of quantized solutions of the wave equation; these solutions are obtained by finding the KAM curves for which the conserved actions are appropriately quantized¹². For general chaotic systems there exists no such correspondence¹². For ARC resonances we propose here a method for establishing an approximate correspondence. It is known for two-dimensional convex ‘billiards’ that the regular orbits which follow KAM curves and generate caustics in the real-space ray motion are well described by an adiabatic approximation¹³. The (adiabatic) KAM curves representing whispering-gallery trajectories in the surface of section of Fig. 2a are given by:

$$\sin \chi(\phi) = \sqrt{1 - (1 - S^2)\kappa(\phi)^{2/3}} \quad (1)$$

where $0 > S > 1$ parametrizes the adiabatic curve and $\kappa(\phi)$ is the curvature of the ARC boundary at an azimuthal angle ϕ . Although for the deformations of interest the relevant adiabatic curves are ‘broken’ and the long-time behaviour of a chaotic whispering-gallery orbit departs strongly from equation (1), we find that for intermediate times (~ 100 – 500 collisions) the orbit still follows closely the nearest adiabatic curve.

Thus we propose that the frequency of the resonance will be well described by the standard semiclassical (eikonal) quantization of the integrable dynamics defined by the adiabatic approximation, even though the resonance lifetime is crucially dependent on the slow chaotic diffusion away from this curve. With this assumption, the resonance frequency can be calculated as a function of deformation by a variant of the eikonal technique of ref. 14; good agreement with exact numerical solutions is found (Fig. 2b). The mode indices (quantized actions) are fixed to their values at zero deformation and the calculation yields both the resonance frequency and the adiabatic curve (value of S) corresponding to a given resonance. This adiabatic curve then defines the initial conditions for our ray calculations of the lifetime of a given resonance.

Using the ray model for the resonance lifetime, we calculate the mean escape time of an ensemble of rays launched with uniform density on the appropriate adiabatic curve. At each collision a ray is allowed to escape with a transmission probability (see Fig. 3 legend) which takes into account both the direct tunnelling which

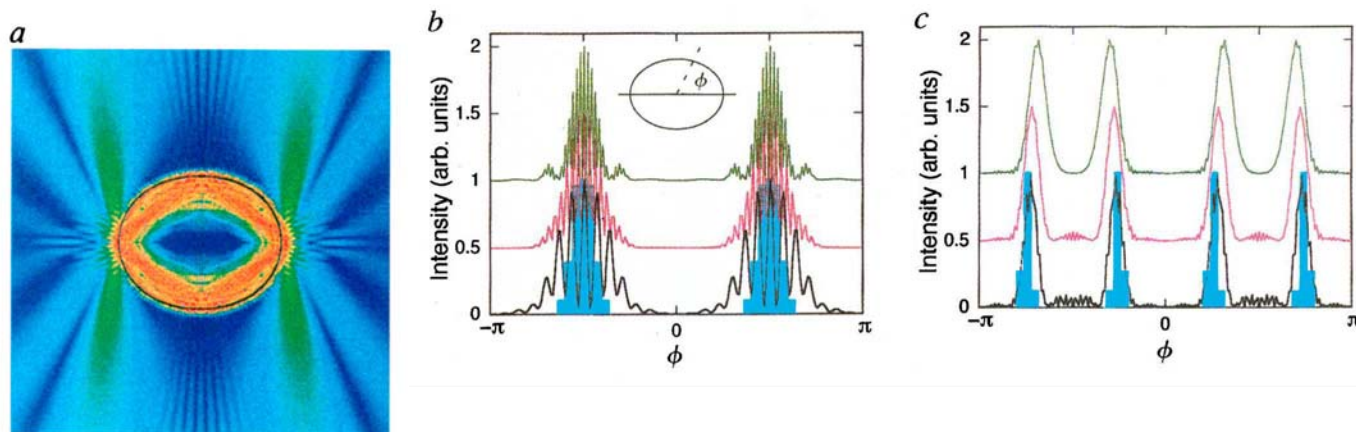


FIG. 4 a, False-colour representation of the electric field intensity in the TM mode for the $kR = 45.14$ resonance at $\epsilon = 0.11$ determined from the exact numerical solution. Intensity is higher for redder colours, and vanishes in the dark-blue regions. Clearly seen are the high-intensity regions in the near field just outside the surface at the highest curvature points $\phi = 0, \pi$, and the high-emission-intensity lines (green) emanating from these points in the tangent directions. b, Far-field plots of the intensity of three resonances with $kR = 12.1$, $\sin \chi \approx 0.8$ (black); $kR = 27.9$, $\sin \chi \approx 0.9$ (red); $kR = 45.4$, $\sin \chi \approx 0.75$ (green). Index of refraction is $n = 2$ (offset

for clarity). Blue histogram is the prediction of the ray-optics model. Note two peaks at $\phi = \pm \pi/2$ corresponding to the tangent directions to the points of highest curvature. c, Far-field plots for three resonances with $kR = 57.8$, $\sin \chi \approx 0.8$ (black); $kR = 48.4$, $\sin \chi \approx 0.91$ (red); $kR = 45.7$, $\sin \chi \approx 0.91$ (green); compared to the ray-optics model (blue histogram) for $n = 1.54$. Note that each peak is split and the intensity is negligible in the direction tangent to the points of highest curvature. This arises from the presence of stable islands eclipsing the points of highest curvature for $n = 1.54$ as shown in Fig. 2a.

is present without deformation, and the Fresnel scattering once $\sin \chi > 1/n$. In Fig. 3 the ray prediction for the lifetime is compared to exact wave solutions for three resonances associated with the same initial adiabatic curves but very different wavelengths. Note that above the critical deformation for refractive escape the widths of all the resonances agree up to factors of the order of unity and also agree well with the ray model. This confirms the universality of the broadening arising from chaotic diffusion.

However, Fig. 3 indicates two significant effects which *do* depend on wavelength and are beyond the ray model. First, we note that for small deformations, where only tunnelling escape is possible, the exact resonance width actually increases much faster than predicted by the ray model. We believe that this is due to chaos-assisted tunnelling^{6,7}, because the direct (angular-momentum-conserving) tunnelling is taken into account by the ray model. All evanescent (tunnelling) processes are forbidden within ray optics, but the probability of such a process decreases exponentially with the tunnelling distance. Consider the whispering-gallery orbit traversing the surface of section of Fig. 2a at $\sin \chi \approx 0.9$. This orbit lies on an unbroken adiabatic curve and hence will never escape according to ray optics. However a wave-packet following this orbit has a chance to tunnel through this dynamical barrier to the edge of the chaotic region at $\sin \chi = 0.8$. From this region, chaotic diffusion will take the wave-packet without any further tunnelling down to the critical angle for refractive escape. Due to the shortness of the tunnelling step, this new escape path, which is introduced by the presence of chaos in a classically inaccessible region of phase space, strongly enhances the evanescent leakage of these resonances.

Second, we find that for resonances with the longest lifetimes the ray model predicts too short a lifetime at large deformations (Fig. 3 inset). At large deformations refractive ray escape is allowed, so the longer lifetime of the wave solution indicates that this classically allowed escape is being suppressed. Precisely such an effect, known as dynamical localization, has been studied extensively in the theory of quantum chaos. Unlike a particle (or ray), a wave is able to explore simultaneously multiple paths while undergoing chaotic diffusion. Typically, after a characteristic time these multiple paths destructively interfere, suppressing further diffusion. Such an effect has been observed indirectly through the suppression of electron ionization in atomic hydrogen in an intense microwave electric field⁸.

We now briefly analyse the onset of directional emission shown in Fig. 4, and discussed in more detail elsewhere¹⁵. As shown in Fig. 2a, the dynamics of whispering-gallery orbits consists of a rapid motion along adiabatic curves and a slow chaotic diffusion transverse to them. Hence the whispering-gallery orbits of interest, which begin far from the critical line $\sin \chi = 1/n$, will diffuse slowly until the adiabatic curve is reached which is tangent to the critical line and then will escape rapidly near the points of tangency (Fig. 2a). These points correspond to the points of maximum curvature of the ARC according to equation (1). Rays escaping near the critical angle are emitted roughly tangent to the surface; this implies strong emission maxima in the far field in directions tangent to the points of maximum curvature (Fig. 4a). This model predicts that all ARC whispering-gallery resonances will have approximately the same directional emission pattern for a given index of refraction (Fig. 4b). This universality persists even when the adiabatic approximation breaks down, as is the case in Fig. 4c where the index of refraction is such that there are islands in the surface of section at the minima of the relevant adiabatic curve. Then the chaotic trajectories circulate around the outside of the islands (Fig. 2a). In this case the peaks split and maximum emission does not occur at the points of highest curvature. All these ray-optics predictions are seen to be in agreement with the wave solutions.

These compact dielectric resonators with controllable Q and highly directional emission may well find applications in micro-lasers and fibre-optic communications. $\square$

Received 12 July; accepted 19 November 1996.

1. Yamamoto, Y. & Slusher, R. E. *Phys. Today* **46**, 66–73 (1993).
2. Nöckel, J. U., Stone, A. D. & Chang, R. K. *Opt. Lett.* **19**, 1693–1695 (1994).
3. Mekis, A., Nöckel, J. U., Chen, G., Stone, A. D. & Chang, R. K. *Phys. Rev. Lett.* **75**, 2682–2685 (1995).
4. Nöckel, J. U. & Stone, A. D. in *Optical Processes in Microcavities* (eds Chang, R. K. & Campillo, A. J.) 389–462 (World Scientific, Singapore, 1996).
5. Johnson, B. R. *J. Opt. Soc. Am.* **10**, 343–352 (1993).
6. Bohigas, O., Boosé, D., Egidio de Carvalho, R. & Marville, V. *Nucl. Phys. A* **560**, 197–214 (1993).
7. Doron, E. & Frisch, S. D. *Phys. Rev. Lett.* **75**, 3661–3664 (1995).
8. Casati, G., Chirikov, B. V., Guarnieri, I. & Shepelyansky, D. L. *Phys. Rev. Lett.* **56**, 2437–2440 (1986).
9. Berry, M. V. *Eur. J. Phys.* **2**, 91–102 (1981).
10. Lazutkin, V. F. *KAM Theory and Semiclassical Approximations to Eigenfunctions* (Springer, Berlin, 1993).
11. Robnik, M. *J. Phys. A* **16**, 3971–3986 (1983).
12. Gutzwiller, M. C. *Chaos in Classical and Quantum Mechanics* (Springer, Berlin, 1990).
13. Robnik, M. & Berry, M. V. *J. Phys. A* **18**, 1361–1378 (1985).
14. Keller, J. B. & Rubinow, S. I. *Ann. Phys.* **9**, 24–75 (1960).
15. Nöckel, J. U., Stone, A. D., Chen, G., Grossman, H. & Chang, R. K. *Opt. Lett.* **21**, 1609–1611 (1996).

ACKNOWLEDGEMENTS. We thank R. Chang, A. Mekis, H. Bruus, G. Hackenbroich and M. Robnik for discussions. This work was supported by the US National Science Foundation and the US Army Research Office.

CORRESPONDENCE should be addressed to A.D.S. (e-mail: stone@venus.cis.yale.edu).

Quantum cryptography on multi-user optical fibre networks

Paul D. Townsend

BT Laboratories, Martlesham Heath, Ipswich IP5 7RE, UK

To establish a secure communication channel, it is necessary to distribute between two users a key which allows safe encryption and decryption of messages. But because decryption is a simple task for any key holder, it is crucial that the key remains secret during distribution. Secrecy cannot be guaranteed if distribution occurs on the basis of classical physical mechanisms, as it is impossible to know whether the key has been intercepted during transmission. Quantum cryptography^{1–3} provides a fundamental solution to this problem. When quantum-mechanical processes are used to establish the key, any eavesdropping during transmission leads to an unavoidable and detectable disturbance in the received key information. Quantum cryptography has been demonstrated using standard telecommunication fibres linking single pairs of users^{4–8}, but practical implementations will require communication networks with many users⁹. Here I introduce a practical scheme for multi-user quantum cryptography, and demonstrate its operation on an optical fibre network. The scheme enables a single controller on the network to establish, and regularly update, a distinct secret key with each network user. These keys can then be used to securely encrypt conventional data transmissions that are broadcast on the network.

The central quantum feature that is exploited in quantum cryptography can be seen from the following *gedanken* experiment, in which a receiver tries to identify the polarization state of an incoming single photon that may be either linearly polarized (vertical $|V\rangle_{\text{lin}}$ or horizontal $|H\rangle_{\text{lin}}$) or circularly polarized (right $|R\rangle_{\text{circ}}$ or left $|L\rangle_{\text{circ}}$). If, for example, the photon is in a $|R\rangle_{\text{circ}}$ state and the receiver happens to try to measure circular polarization an accurate deterministic result will be obtained, that is, $|R\rangle_{\text{circ}}$. However, if the receiver tries a linear measurement the outcome will be unpredictable, yielding the result $|V\rangle_{\text{lin}}$ with 50% probability or $|H\rangle_{\text{lin}}$ with 50% probability. Furthermore, after the measurement the photon is repolarized into a linear state and the receiver cannot reconstruct the original $|R\rangle_{\text{circ}}$ state or even know that an inaccurate outcome has occurred. These scenarios reflect the fact that linear polarization and circular polarization are complementary quantum observables that are incompatible in

the sense that a measurement of one property will always disturb the other¹⁰.

In quantum cryptography these properties are exploited for secure key sharing between a single transmitter-and-receiver pair conventionally referred to as 'Alice' and 'Bob', whose communication may be intercepted by an eavesdropper 'Eve'. In this scheme^{1,2} quantum information is transmitted from Alice to Bob using the single photon polarization states to represent the boolean values 0 and 1 (that is, $|V\rangle_{\text{lin}} \equiv 0$, $|H\rangle_{\text{lin}} \equiv 1$ and $|R\rangle_{\text{circ}} \equiv 0$, $|L\rangle_{\text{circ}} \equiv 1$). Key distribution begins with Alice transmitting to Bob a timed sequence of photons, each of which is randomly modulated with one of the four possible polarization states. This is equivalent to Alice sending a random bit sequence in which she randomly chooses between the circular or linear representation for each bit. Bob attempts to measure a photon in each time period, making a random and independent choice of either linear or circular for his polarization measurement. If he detects a photon (possibly a rare event if there is loss in the system) he records the time period and the result of the measurement, for example, $|L\rangle_{\text{circ}} = 1$. At the end of this process, Alice and Bob use a conventional (insecure) classical communication link to carry out a public discussion about the quantum transmission. Bob reveals to Alice the time periods when he received a photon and the type of measurement that he performed in each (for example, linear) but not the result (that is, $|V\rangle_{\text{lin}}$ or $|H\rangle_{\text{lin}}$) because this would give away the bit. Alice knows what she sent in each of these time periods, and she instructs Bob to discard all the data from the instances where he performed an incompatible measurement leading to a probabilistic outcome. At this stage Bob is in possession of a 'raw' random bit string that has been 'distilled' from Alice's initial random bit sequence. If Eve wants to intercept the transmission she must perform measurements on the sequence of photons. However, because she has no *a priori* knowledge of the coding representation chosen by Alice for each photon, her measurements will frequently be incompatible with the current representation leading to a probabilistic outcome. Without the benefit of the post-transmission public discussion Eve is unable to identify these instances, and the sequence of polarized photons

that she re-sends on to Bob will be corrupted. Consequently, if Alice and Bob perform a classical error detection and correction technique (using the public channel) to assess the bit rate error rate ratio (BER) for the raw bit string, they have a fundamental measure of the level of eavesdropping^{2,11,12}. If this is below a certain bound¹³ then a classical data manipulation technique called privacy amplification^{14,15} can be applied to generate a new shared random bit string about which Eve has only an arbitrarily small amount of information. This final bit string is now certifiably secret and is safe to use as a cryptographic key.

Quantum cryptography can also be applied to the case of a passive optical network (PON) comprising a single network controller, Alice, linked via a sequence of passive optical splitters to a multiple of network users, Bob₁–Bob_N. In the classical limit, information transmitted by Alice in the downstream direction on the PON is broadcast to all Bobs; similarly, upstream transmissions by any of the Bobs are 'broadgathered' by Alice. If Alice can establish a verifiably secret and unique individual key with each Bob, she can then use (for example) the *i*th key to encrypt information intended for Bob_i. In this way the network can be operated securely because, although the encrypted information is broadcast on the network, Alice and Bob_i can be confident that no other network user or external eavesdropper has obtained any information on their shared key. Quantum cryptography can be used to perform the key distribution task simply by exploiting the quantum level behaviour of the optical splitters in the PON. A single photon incident on such a node cannot split, and is instead

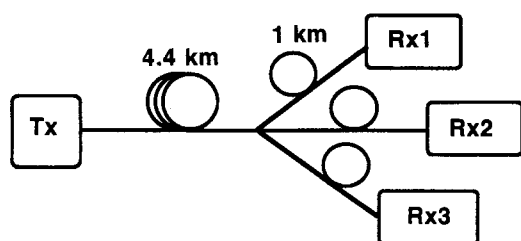


FIG. 1 Experimental passive optical network based on standard telecommunications fibre. The transmitter (Tx) was linked to each receiver (Rx1–Rx3) via a 4.4-km length of fibre (triple circle), a 1×3 splitter, and a 1-km length of fibre (single circle), giving an effective link length of 5.4 km. The optical source in the transmitter was an externally modulated 1.3- μm -wavelength semiconductor laser which generated 200-ps-duration pulses at a repetition rate of 1 MHz. This was followed by a variable attenuator that controlled the mean photon number μ of the pulses at the network input. The quantum coding scheme was based on an interferometric scheme implemented using lithium niobate phase modulators in the transmitter and receivers. Detection was carried out using cooled germanium avalanche photodiodes (APDs) operating in a time-correlated photon-counting mode^{5,18}. The optical transmitter and receivers were controlled by two personal computers that were also linked via an electronic network that covers the BT Laboratories site. This electronic network was used for the public discussion phase of the protocol, and also for the transmission of data encrypted using the secret keys distributed over the PON. All of this traffic could also be carried by the PON by using wavelength division multiplexing to separate the quantum key distribution channel from a conventional high-bandwidth channel operating at, for example, 1.55 μm wavelength on the same network.

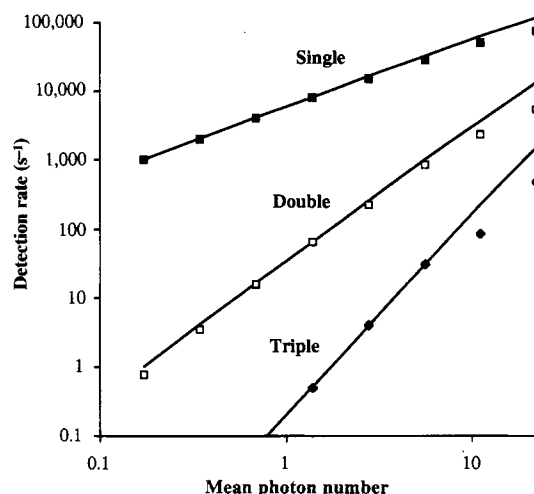


FIG. 2 Effective detection rates for successful pulses (that is, after approximately half of the results have been discarded during public discussion) plotted against μ , the mean photon number per pulse at the network input. Solid squares, pulses detected at any single receiver; open squares, pulses detected at any two of the receivers; filled diamonds, pulses detected at all three receivers. Detection events or counts are identified with a particular input pulse by means of the time period in which they occurred, or combinations of receivers, on the network. The data can be modelled by assuming a poissonian source and detectors that only distinguish between pulses containing 0 photons and pulses containing 1 or more photons. In this case the successful pulse rates r_x for single ($x = 1$), double-shared ($x = 2$) and triple-shared ($x = 3$) pulses are given by $r_x \approx f \{ \frac{1}{2} [1 - \exp(-\frac{1}{2} \alpha \eta \mu)] \}^x$, where f is the laser pulse rate, α is the transmission fraction after loss and η is the detector quantum efficiency. The factor of 1/2 is due to the fact that approximately half of the detection events are discarded after the initial public discussion phase, and the factor of 1/3 in the exponential represents the splitting factor for the 1×3 fibre coupler. The solid lines show the predictions of the equation using the experimentally measured values of $f = 10^6$ Hz, $\alpha = 0.16$ and $\eta = 0.22$. The results are in close agreement with the experimental points except at high μ values, where the onset of detector saturation tends to reduce the observed count rates.

routed to one (and only one) of the output paths. The choice of output path for any particular photon is random and unpredictable, governed only probabilistically by the classical splitting ratios for each path. Consequently, if the conventional scheme discussed above is applied to the network, then each Bob will be supplied with a unique randomly selected subset of bits from the sequence that Alice transmits into the PON. By carrying out the post-transmission public discussion with each Bob in turn, Alice can identify which photons were shared with each receiver, and (via the error rate) guarantee secrecy from external eavesdroppers.

These ideas were tested using the simple 1×3 PON shown in Fig. 1. The experiment had two main differences from the polarization example discussed above. First, because weak coherent pulses of mean photon number $\mu < 1$ behave very much like single photons at the network splitters, but are much easier to generate in practice, a conventional pulsed semiconductor laser was used for the optical source. Second, the system employed an interferometric encoding scheme^{4,5,16} directly analogous to the polarization technique but with the linear and circular states replaced by four phase-shift values. The latter were applied using phase modulators in the transmitter and receivers. Figure 2 shows the results of an experiment designed to measure the splitting probability for the 'successful' pulses, that is, those pulses that are kept after the first stage of public discussion when the probabilistic outcomes are discarded (but before error correction and privacy amplification). As these pulses generate the raw random bit strings, a low splitting probability demonstrates a low degree of correlation between the three individual bit strings. The results show a linear reduction in the detection rate of successful pulses at any single receiver as the mean photon number μ decreases. However, the rates for shared successful pulses fall much more rapidly with μ , varying quadratically for pulses shared by any two receivers, and cubically for pulses shared by all three receivers. Consequently, at the lowest value of $\mu = 0.18$ (Fig. 2), any individual raw random bit string contains (on average) only about 1 in 10^3 bits that are shared with one other receiver, and only about 1 in 10^6 bits that are shared with both of the other receivers on the network. This level of correlation is

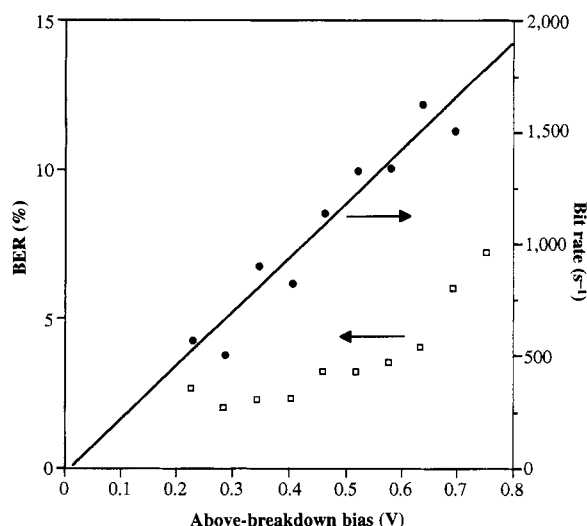


FIG. 3 Mean values for the bit rates (filled circles) and bit rate error rate ratios (BERs; open squares) obtained for the three receivers on the network during key distribution. The results are plotted as a function of the above-breakdown APD bias V_b . The quantum efficiency of the detectors varied approximately linearly with V_b , as illustrated by the linear fit to the bit rate data, so the bit rate in the system could be increased simply by increasing the bias voltage. However, because the error rate rises sharply above $V_b \approx 0.65$ V, due to an increase in dark counts in the APDs¹⁸, it was preferable to restrict operation to $V_b \leq 0.6$ V.

extremely low, and demonstrates that the low-intensity pulses are behaving effectively like indivisible particles that are randomly routed by the network splitter. Further experiments were carried out at a μ value of 0.18 to monitor the effective bit rates and BERs for the raw random bit strings. The results are shown in Fig. 3, plotted as a function of the above-breakdown bias voltage applied to the avalanche photodiode (APD) detectors. Bit rates of ~ 1 kbit s^{-1} and BERs of $\sim 3\%$ were achieved by all three receivers on the PON. Although the measured BERs were dominated by noise due to APD dark counts, a worst-case assumption was made that the disturbance was entirely due to eavesdropping². In this case, the upper bound on the amount of information leakage was low enough that error correction and privacy amplification could be used to generate the final highly secret keys. For comparison, recent calculations¹³ suggest that secure key distribution is possible for BERs up to $\sim 15\%$, so the experimental BERs $\sim 3\%$ were comfortably below this threshold.

The feasibility of quantum key distribution on optical fibre networks has been demonstrated, but how practical are such communication schemes? Considering first the performance of the present system: I note that as most modern encryption schemes utilize keys of the order of 1 kbit or less, the effective key distribution rate of ~ 1 kbit s^{-1} attained experimentally may be more than adequate for many potential applications. Furthermore, because the system is based on standard telecommunications components and operates at a wavelength of $1.3 \mu m$, it is compatible with installed fibre infrastructures. On the negative side, the present system has only relatively modest values of link length $L = 5.4$ km and splitting factor $N = 3$. If either of these parameters were increased, the photon count rate at the detectors would fall and, as the dark count rate for the APDs remains constant, the BER would rise. For example, if L and N were increased to 13 km and 8 respectively, with the present germanium APDs, the BER would reach the threshold value of 15% and secure key distribution would become impossible. But these limits are not fundamental, and system performance will improve as single photon detector technology for $1.3 \mu m$ and $1.5 \mu m$ telecommunication wavelengths continues to mature. Another approach is to use the wavelength range around $0.8 \mu m$ for key distribution. This allows the use of silicon APDs that are currently at a more advanced stage of development, and which offer dark count rates more than 10^2 times lower than germanium devices¹⁷. If such detectors were used in the present system, for example, a similar link length and BER performance could be obtained with a much larger splitting ratio of $N \approx 50$. Hence applications in secure local area networks appear, at least technically, quite feasible. $\square$

Received 2 August; accepted 19 November 1996.

1. Bennett, C. H., Brassard, G., Breidbart, S. & Wiesner, S. *Advances in Cryptology: Proceedings of Crypto '82* (eds Chaum, D., Rivest, R. L. & Sherman, A. T.) 267–275 (Plenum, New York, 1983).
2. Bennett, C. H., Bessette, F., Brassard, G., Salvail, L. & Smolin, J. J. *J. Cryptol.* **5**, 3–28 (1992).
3. Ekert, A. K. *Phys. Rev. Lett.* **67**, 661–663 (1992).
4. Townsend, P. D. *Electron. Lett.* **30**, 809–811 (1994).
5. Marand, C. & Townsend, P. *Opt. Lett.* **20**, 1695–1697 (1995).
6. Franson, J. D. & Jacobs, B. C. *Electron. Lett.* **31**, 232–234 (1995).
7. Hughes, R. J., Luther, G. G., Morgan, G. I. & Simmons, C. in *Proc. 7th Rochester Conf. on Coherence and Quantum Optics* (eds Eberly, J. H., Mandel, L. & Wolf, E.) 103–112 (Plenum, New York, 1996).
8. Müller, A., Zbinden, H. & Gisin, N. *Nature* **378**, 449 (1995).
9. Townsend, P. D., Phoenix, S. J. D., Blow, K. J. & Barnett, S. M. *Electron. Lett.* **30**, 1875–1876 (1994).
10. Phoenix, S. J. D. & Townsend, P. D. *Contemp. Phys.* **36**, 165–195 (1995).
11. Ekert, A. K., Huttner, B., Palma, G. M. & Peres, A. *Phys. Rev. A* **50**, 1047–1056 (1994).
12. Huttner, B. & Ekert, A. K. *J. Mod. Opt.* **41**, 2455–2466 (1994).
13. Lütkenhaus, N. *Phys. Rev. A* **54**, 97–111 (1996).
14. Bennett, C. H., Brassard, G. & Robert, J.-M. *Soc. Ind. Appl. Math. J. Comp.* **17**, 210–229 (1988).
15. Bennett, C. H., Brassard, G., Crépeau, C. & Maurer, U. M. *IEEE Trans. Inform. Theory* **41**, 1915–1923 (1995).
16. Bennett, C. H. *Phys. Rev. Lett.* **68**, 3121–3124 (1992).
17. Townsend, P. D., Marand, C., Phoenix, S. J. D., Blow, K. J. & Barnett, S. M. *Phil. Trans. R. Soc. Lond. A* **354**, 805–817 (1996).
18. Owens, P. C. M., Rarity, J. G., Tapster, P. R., Knight, D. & Townsend, P. D. *Appl. Opt.* **33**, 6895–6901 (1994).

ACKNOWLEDGEMENTS. I thank K. Blow and S. Phoenix for many illuminating discussions during the course of this work.

Acceleration of a Diels–Alder reaction by a self-assembled molecular capsule

Jongmin Kang & Julius Rebek Jr

Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA and The Skaggs Institute for Chemical Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA

THE interior of cage-like molecules can be considered to provide a new phase of matter^{1,2}, in which it becomes possible to stabilize reactive intermediates³ and to observe new forms of stereoisomerism⁴. Cage-like molecular complexes that self-assemble through weak intermolecular forces are dynamic species^{5,6}, encapsulating guest molecules reversibly. They can persist over timescales ranging from microseconds to hours, long enough for chemical processes to take place within them. Here we report the acceleration of a Diels–Alder reaction by encapsulation of the reactants in a self-assembling molecular capsule^{7,8}. Although product inhibition (lack of dissociation) prevents the system from showing true catalytic behaviour, there is clear evidence for a rate increase of over two orders of magnitude owing to the effective enhancement of concentration inside the capsule.

Compound **1** (Fig. 1) and its congeners self-assemble in organic solvents to form dimeric capsules^{7,8}. The molecules feature self-complementary patterns of hydrogen-bonding sites and a dimer is formed when the concave surfaces of two molecules come together (Fig. 1). Intermolecular hydrogen bonds hold the two subunits together in much the same manner that the stitches along the seam hold a softball together. In aromatic solvents such as benzene, the dimerization constant is large ($>10^6 \text{ M}^{-1}$) and the dynamics of assembly are slow on the NMR timescale but fast on the human timescale ($< \text{seconds}$). The dimers can encapsulate molecules of complementary size and shape and do so reversibly. Adamantanes, ferrocene derivatives and [2,2] paracyclophane are among the sizeable guests that fit well within these capsules. The demonstration that two molecules of solvent benzene are accommodated inside⁸, raised the possibility of the use of these capsules as chambers for bimolecular reactions.

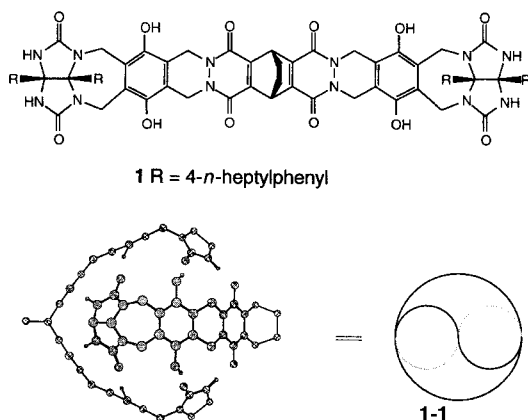
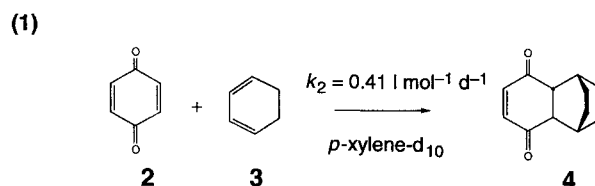


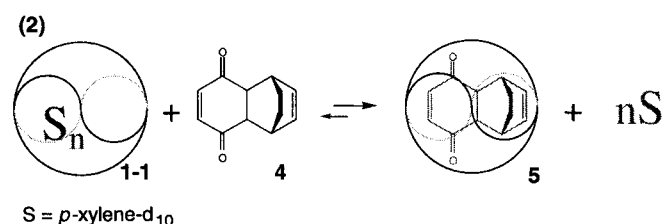
FIG. 1 Representations of the monomeric **1** and its self-assembling dimeric form **1-1**. The hydrogen-bonding network of the dimeric form as an energy-minimized structure generated by MacroModel²² is highlighted in the cross-section on the left; other atoms have been deleted for viewing clarity. The 'softball' representation is on the right.

With this in mind, we investigated the Diels–Alder reaction between *p*-quinone **2** and cyclohexadiene **3**:



The reaction is slow at room temperature. In *p*-xylene-*d*₁₀ at molar concentrations of each component, the half-life is about two days ($k_2 = 0.41 \text{ l mol}^{-1} \text{ d}^{-1}$); at millimolar concentrations no product is detected by NMR in a week, as the corresponding half-life is of the order of a year. The stereochemistry of the product **4** is that anticipated for the usual *endo* orientation of reactants in the transition state.

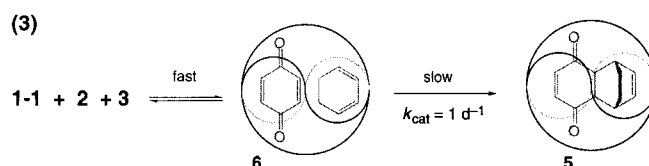
In dilute solution (for example, 1 mM) the Diels–Alder adduct **4** is readily encapsulated by the dimeric **1-1**:



and NMR signals unique to the encapsulated species **5** can be observed. The widely separated and relatively sharp signals for the free and bound forms of **4** indicate slow exchange of the guest into and out of the host capsule on the NMR timescale (Fig. 2*a–c*). The adduct is a welcome guest; the binding affinity is too large for accurate determination through NMR titrations, and only a lower limit, $K_a > 10^5 \text{ M}^{-1}$ can be placed on the association constant. This figure leaves little doubt that the cavity is roomy enough to accommodate the transition-state geometry of this Diels–Alder reaction.

With both *p*-quinone and cyclohexadiene (4 mM each) are added to a solution of the dimeric capsule (1 mM) in *p*-xylene-*d*₁₀ at room temperature, a new, well defined complex emerges, as evidenced by the NMR spectrum (Fig. 3*a, b*). One species dominates in the spectrum and a signal for encapsulated quinone is observed. No signals unique to encapsulated cyclohexadiene can be assigned, but some must be present inside because the signals for the encapsulated adduct appear within one day. These signals continue to grow at the expense of the signal for encapsulated quinone (Fig. 3*c, d*).

The probably precursor to **5** (complex **6**, reaction (3)) can be regarded as the counterpart of a Michaelis complex, as it converts to product in a first-order process. Consistent with this treatment, observed rates at various cyclohexadiene concentrations (1–7 mM) show saturation kinetics; the parameters at ambient temperature are $V_{\text{max}} = 0.001 \text{ M}^{-1} \text{ d}^{-1}$ and $k_{\text{cat}} = 1 \text{ d}^{-1}$ (Fig. 4). The implication is that the Diels–Alder reaction takes place inside the capsule:



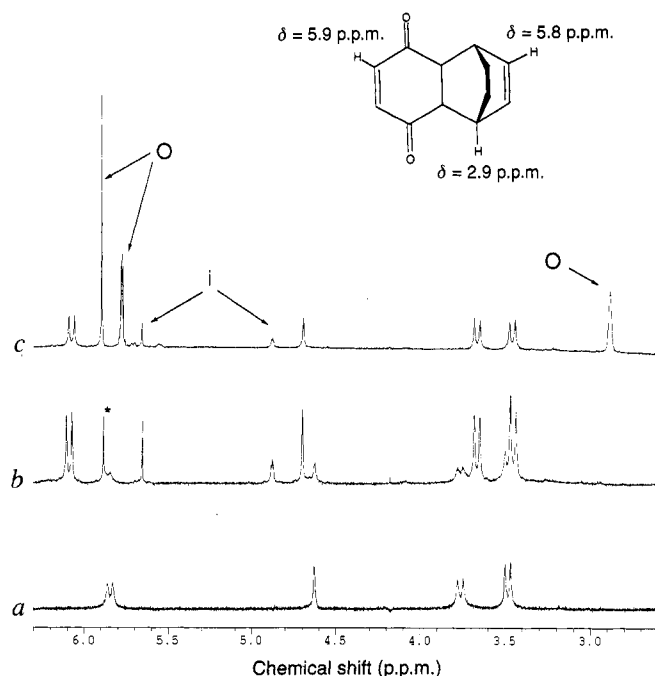


FIG. 2 Proton NMR spectra of **1-1** and its encapsulation complex with the adduct **4** in *p*-xylene- d_{10} . Signals of the guest inside the capsule and outside (free) are labelled with 'i' and 'o', respectively. The signal designated with an asterisk represents chloroform. a, Compound **1-1** alone. b, Compound **1-1** with 0.7 equiv. of Diels–Alder adduct **4** added; all of the latter is encapsulated and separate signals are observed for the complex **5** and solvated dimer **1-1**. c, Compound **1-1** with 6 equiv. of Diels–Alder adduct **4** added; all of the capsule is occupied as complex **5** and separate signals are observed for free and encapsulated adduct **4**. Assignments for resonances of free **4** are shown on the structure.

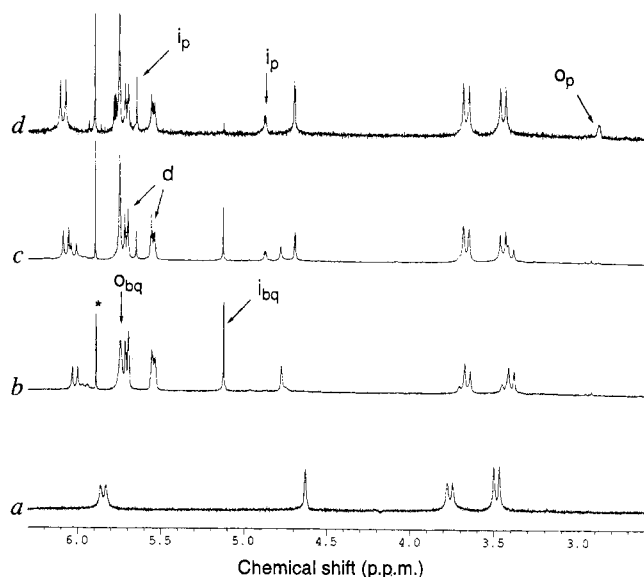


FIG. 3 Changes in the proton NMR spectra of **1-1** during the reaction of *p*-benzoquinone **2** and 1,3-cyclohexadiene **3** in *p*-xylene- d_{10} . The signals for *p*-benzoquinone (bq) and Diels–Alder product (p) inside and outside are designated as 'i' and 'o' respectively. Signals from cyclohexadiene are designated as 'd'. The signal designated with an asterisk represents chloroform impurity. a, Compound **1-1** alone. b, Shortly after 4 equiv. of *p*-benzoquinone and 4 equiv. of 1,3-cyclohexadiene were added to the solution of **1-1**. c, The reaction mixture after 2 days. d, The reaction mixture after 19 days, showing evidence of released product and **5** as the principal species.

After several days at ambient temperature some of the adduct appears outside the capsule (Fig. 3d), but it is clear that product inhibition is a serious problem and prevents the system from turning over and being an effective catalyst. Indeed, addition of adduct **4** or other good guests (benzene, [2,2] paracyclophane) effectively inhibit the reaction. Nonetheless, the effect of the dimer **1-1** on the reaction rate is impressive; the reaction is accelerated some 200-fold under these conditions.

Are the rate accelerations provided by **1-1** for the Diels–Alder reaction due to encapsulation of the components, or some more conventional form of catalysis? For example, the Diels–Alder reaction is known to be catalyzed by biphenylenediols⁹ that can form hydrogen bonds with the carbonyl oxygens of suitable dienophiles. Many hydrogen-bonding sites are featured on the periphery of **1**, but two control experiments to test this possibility with **9** and **10** (Fig. 5) argue against this form of catalysis. The first involved a molecule that bears the same functionalities as **1** but differs in shape, while the second involved a molecule that bears the same shape as **1** but differs in functionalities. Neither **9** nor **10**

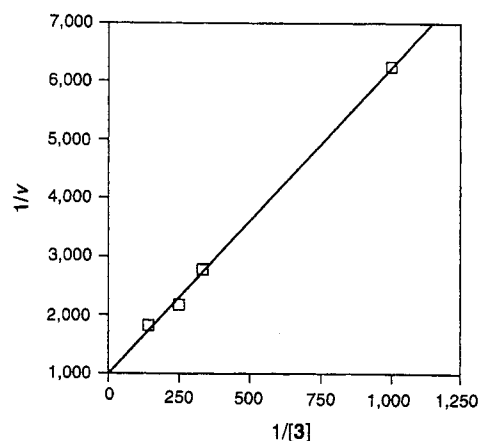


FIG. 4 Lineweaver–Burke plot of the reaction rates, V , of **2** with **3** in the presence of **1-1**. The concentrations of 1,3-cyclohexadiene **3** were changed with a constant concentration (4 mM) of *p*-benzoquinone.

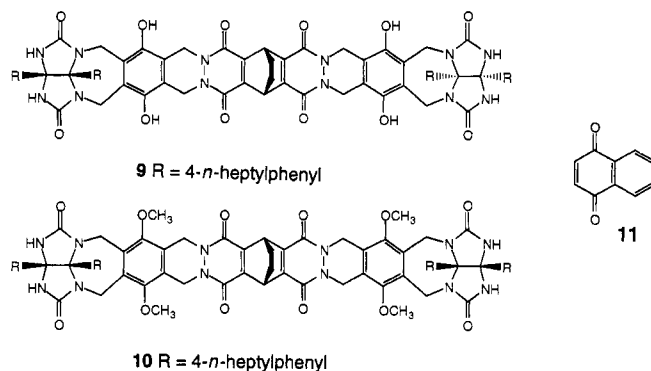


FIG. 5 Compounds used as controls. The reaction of **2** with **3** was not accelerated by the S-shaped stereoisomer **9** in which the glycoluril substituents are on opposite faces of the structure. Nor was it accelerated by the C-shaped structure **10** in which the phenolic units are methylated and the glycoluril substituents are on the same face of the structure. The dimeric form **1-1** did not effect the rate of the Diels–Alder reactions between naphthoquinone **11** and **3**.

accelerated the reaction. Accordingly, acid catalysis through hydrogen bonding to the phenols or the glycoluril functions can be excluded as the source of acceleration. Although some form of acid catalysis might operate within the special microenvironment presented by the capsule to guests, this possibility remains untested, and may even be untestable. Finally, the corresponding reaction of 1,4 naphthoquinone **11**, a molecule too large for the capsule, was not accelerated by **1-1**. The factors of size selectivity, saturation kinetics, and product inhibition all support an encapsulated transition state.

The transition-state geometry of this reaction leading to the *endo* product is reminiscent, in size and shape, of two encapsulated benzene molecules oriented in a face-to-face stacking interaction. Accordingly, the acceleration of the Diels–Alder reaction inside the capsule is not unreasonable. Acceleration of the reverse of the Diels–Alder reaction by encapsulation is unreasonable, and we interpret the accelerating effect of encapsulation on the forward rate as resulting from an enhanced concentration of the reactants inside the capsule, rather than some special stabilization of the transition state. Using a technique described elsewhere¹⁰, we estimate the interior volume of the capsule to be $\sim 300 \text{ \AA}^3$. When this space is occupied with one molecule of quinone and cyclohexadiene, the concentration of each reactant inside the capsule is calculated to be about 5 M—this is 1,000 times the concentration of the reactants in the bulk solution (4 mM). Consistent with these estimates, the effective molarity¹¹ calculated from the rate data ($k_{\text{cat}}/k_{\text{uncat}}$) gives a value of 2.4 M for the encapsulated reactants.

The results here augur well for the application of reversibly formed molecular assemblies as reaction chambers. If the size and shape selectivity observed in their binding of ground states can be extended to preferential recognition of transition states^{12,13} or even high-energy intermediates, then true catalysis via encapsulation could result. To release the product, it may be possible to use the approach devised by Hilvert¹⁴ in antibody catalysis, which takes advantage of a subsequent reaction of the Diels–Alder product that alters its shape and affinity for the binding site and forces turnover. The general problem of product inhibition may be overcome in the future by using dissociative processes, reactions that result in an increased number of encapsulated species. At present, the evidence reported here indicates that in addition to heat, pressure, acids, antibodies¹⁴, micelles^{15–18}, medium^{19,20} and template effects²¹, Diels–Alder reactions can also be promoted by reversible encapsulation. □

Infrared remote sensing of breaking waves

A. T. Jessup*, C. J. Zappa*, M. R. Loewen† & V. Hesany*

* Applied Physics Laboratory, University of Washington, Seattle, Washington 98195, USA

† Department of Mechanical and Industrial Engineering, University of Toronto, Toronto, Ontario, M5S 3G8, Canada

ENERGY dissipation due to deep-water wave breaking plays a critical role in the development and evolution of the ocean surface wave field. Furthermore, the energy lost by the wave field via the breaking process is a source for turbulent mixing and air entrainment, which enhance air–sea heat and gas transfer^{1–3}. But the current lack of reliable methods for measuring energy dissipation associated with wave breaking inhibits the quantitative study of processes occurring at ocean surfaces, and represents a major impediment to the improvement of global wave-prediction models⁴. Here we present a method for remotely quantifying wave-breaking dynamics which uses an infrared imager to measure the temperature changes associated with the disruption and recovery of the surface thermal boundary layer (skin layer). Although our present results focus on quantifying energy dissipation—in particular, we show that the recovery rate of the skin layer in the wakes of breaking waves is correlated with the energy dissipation rate—future applications of this technique should help to elucidate the nature of important small-scale surface processes contributing to air–sea heat⁵ and gas⁶ flux, and lead to a fuller understanding of general ocean–atmosphere interactions.

Under most circumstances, a net upward heat flux from the ocean to the atmosphere occurs by molecular conduction through the skin layer at the ocean surface. As a result, the surface, or skin, temperature of the ocean is less than the bulk temperature immediately below by a few tenths of a degree Celsius^{7,8}. An infrared radiometer measures only the skin temperature because the optical depth of the infrared radiation detected, about 10 μm , is much less than the thickness of the skin layer, of the order of 1 mm (refs 9–11). Our measurements confirm an early result¹² showing that when the cool skin layer is momentarily disrupted by a breaking wave, the skin temperature of the resulting turbulent wake is approximately equal to the bulk temperature. As the wake subsides, the skin layer recovers, and the skin temperature returns to its original, cooler value. Infrared measurements of breaking waves made in the surf zone¹² show a recovery time of the order of 10 s. We have found that the time needed for the skin layer to recover depends on the ambient heat flux¹³ and, most significantly for quantifying wave breaking, on the strength of the event. Therefore, infrared measurements of the skin-layer recovery process should provide the capability to remotely monitor free-surface turbulence under conditions of constant heat flux or when the heat flux dependence is known.

Our measurements were made using an infrared imager which provides a time series of two-dimensional images of the skin temperature as inferred from the infrared radiance. Figure 1 shows a sequence of simultaneous video and infrared images of a breaking wave that was taken in the open ocean (time increases *a* to *d*). The first three video images show the breaking wave propagating from right to left as it evolves from a narrow, curved ridge to an actively breaking crest, or whitecap. The first three infrared images show features of increased temperature that clearly correspond to the whitecap, which can be up to 0.1 °C warmer than its wake. Part of the whitecap signature is probably an apparent temperature increase due to increased emissivity¹⁴ associated with surface roughness and bubbles. Of primary interest here is the portion of the wake where the temperature changes are associated with the skin-layer disruption due to breaking-

Received 14 August; accepted 31 October 1996.

- Sherman, J. C. & Cram, D. J. *J. Am. Chem. Soc.* **117**, 4527–4528 (1989).
- Garel, L., Dutasta, J.-P. & Collet, A. *Angew. Chem. Int. Edn Engl.* **32**, 1169–1171 (1993).
- Cram, D. J., Tanner, M. E. & Thomas, R. *Angew. Chem. Int. Edn Engl.* **30**, 1024–1027 (1991).
- Timmerman, P., Verboom, W., van Veggel, F. C. J. M., van Duynhoven, J. P. M. & Reinhoudt, D. N. *Angew. Chem. Int. Edn Engl.* **33**, 2345–2348 (1994).
- Wyler, R., de Mendoza, J. & Rebek, J. *Jr Angew. Chem. Int. Edn Engl.* **32**, 1699–1701 (1993).
- Valdés, C., Spitz, U. P., Toledo, L., Kubik, S. & Rebek, J. *Jr J. Am. Chem. Soc.* **117**, 12733–12745 (1995).
- Meissner, R. S., de Mendoza, J. & Rebek, J. *Jr Science* **270**, 1485–1488 (1995).
- Kang, J. & Rebek, J. *Jr Nature* **382**, 239–241 (1996).
- Kelly, T. R., Meghani, P. & Ekkundi, V. S. *Tetrahed. Lett.* **31**, 3381–3384 (1990).
- Meissner, R., Garcias, X., Mecozzi, S. & Rebek, J. *Jr J. Am. Chem. Soc.* (in the press).
- Kirby, A. J. *Adv. Phys. Org. Chem.* **17**, 183–278 (1980).
- Mock, W. L., Irra, T. A., Wepsiec, J. P. & Adhya, M. J. *Org. Chem.* **54**, 5302–5308 (1989).
- McCurdy, A., Jimenez, L., Stauffer, D. A. & Dougherty, D. A. *J. Am. Chem. Soc.* **114**, 10314–10321 (1992).
- Hilvert, D., Hill, K. W. & Auditor, M. T. M. *J. Am. Chem. Soc.* **111**, 9261–9263 (1989).
- Grieco, P. A., Garner, P. & He, Z.-M. *Tetrahed. Lett.* **24**, 1897–1900 (1983).
- Breslow, R., Maitra, U. & Rideout, D. *Tetrahed. Lett.* **24**, 1901–1904 (1983).
- Braun, R., Schuster, F. & Sauer, J. *Tetrahed. Lett.* **27**, 1285–1288 (1986).
- Singh, V. K., Raju, B. N. S. & Deota, P. T. *Synth. Commun.* **18**, 567–574 (1988).
- Grieco, P. A., Kaufman, M. D., Daebule, J. F. & Saito, N. J. *Am. Chem. Soc.* **118**, 2095–2096 (1996).
- Breslow, R. & Guo, T. J. *J. Am. Chem. Soc.* **110**, 5613–5617 (1988).
- Walter, C. J., Anderson, H. L. & Sanders, J. K. M. *J. Chem. Soc. Chem. Commun.* 459–460 (1993).
- Mohamadi, F. et al. *J. Comput. Chem.* **11**, 440–467 (1990).

ACKNOWLEDGEMENTS. We thank J. Brauman, D. Hilvert, J. Stubbe, J. Williamson and M. Heagy for discussions. This work was supported by the National Institutes of Health.

CORRESPONDENCE should be addressed to J.R. at the Skaggs Institute for Chemical Biology (e-mail: jrebek@scripps.edu).

generated turbulence. In the fourth infrared image, a roughly circular region of increased temperature extends significantly beyond the area covered by the whitecap and residual foam in the corresponding video image. The extent and location of this warm patch imply that it corresponds to the turbulent wake of the breaking wave.

The crest speed of a breaking wave has been suggested as a useful measure of its scale¹⁵ and has been used as such in field measurements of the breaking process¹⁶. We have found that the time required for the skin layer within the wake to recover is proportional to the speed of the breaking crest. Faster-moving waves dissipate more energy and produce more turbulence, resulting in longer skin-layer recovery times. Figure 2a shows time series of the normalized skin temperature deviation, T_n , at a location within the turbulent wake for two breaking waves with different crest speeds, C_b , and correspondingly different recovery times, τ . The rapid increase in T_n due to the passage of the whitecap is followed by a gradual decrease which takes longer for the faster-moving wave. In Fig. 2b, τ is plotted versus C_b for 36 events observed during a 6-hour period of steady environmental conditions (see Fig. 2 legend for details). Although the results shown in

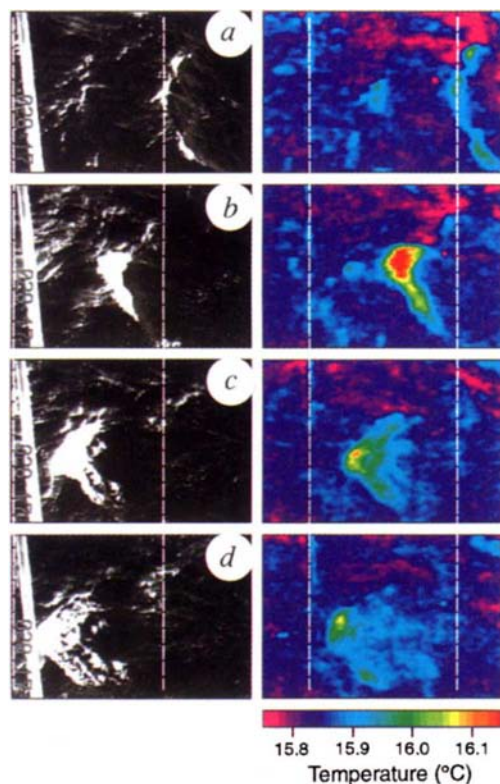


FIG. 1 Sequence of simultaneous, co-located video images (left) and infrared images (right) of a breaking wave in the open ocean. Image size is approximately 5 m $\times$ 10 m. The breaking wave is propagating from right to left; time increases down the page and spans 1.6 s. For reference purposes, each image includes two dashed vertical lines corresponding to roughly the same locations; the right-hand line denotes the leading edge of the breaking wave crest in the first frame. The whitecap in the video images appears as the warmest region in the infrared images. After the crest spills, a roughly circular warm patch remains (bottom-right image) which is significantly removed from the area dominated by the whitecap and foam. The infrared signature of the turbulent wake results from the disruption of the cool skin layer by the breaking wave. For this example, the bulk-skin temperature difference was $\sim 0.15^\circ\text{C}$ and the wind speed was 7.7 m s^{-1} . The measurements were made using an infrared imager (Agema Model 880LW) operating at wavelengths of 8 to $12\text{ }\mu\text{m}$ aboard the research platform RP FLIP off San Diego, California. The temperature differences between the wake and undisturbed regions of the ocean surface that were derived from the infrared imagery were consistent with independent measurements of the bulk-skin temperature difference.

Fig. 2b demonstrate that the recovery time is correlated with crest speed, this scaling alone is not sufficient to conclude that the skin-layer recovery process can be used to quantify the dynamics of wave breaking.

To link the skin-layer recovery directly to the energy loss due to wave breaking, we examine the correlation between a measure of the skin-layer recovery rate, R_1 , and an estimate of the energy dissipation rate per unit area, E . From the time series of skin temperature within the wake, the recovery rate is estimated by $R_1 = \Delta T/\tau$, where ΔT is the temperature deviation corresponding to the change in T_n from its maximum to a level of 0.3 (see Fig. 2a). A measure of R_1 that is dimensionally consistent with E (W m^{-2}) is given by the rate of heat transfer per unit area required to re-establish the skin layer; we define this parameter as the restoring heat flux, $Q_r = (\rho c_p \delta) R_1$, where ρ is the density of water, c_p is the

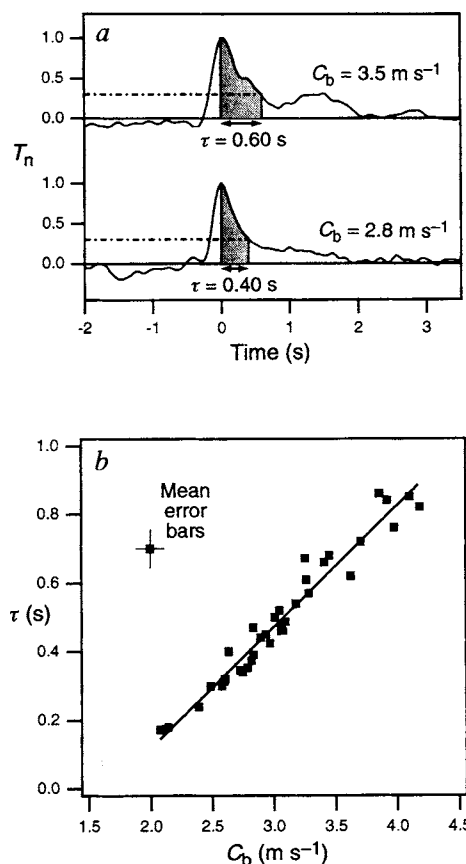


FIG. 2 a, Time series of T_n , the skin temperature deviation normalized by its maximum, for two oceanic breaking waves with different crest speeds, C_b , and skin-layer recovery times, τ . The skin temperature increases rapidly as the whitecap passes and then decays as the skin layer recovers. A three-step procedure was applied to infrared sequences like those in Fig. 1 to determine the crest length, crest speed, and recovery time. First, the area in each image was determined by applying a threshold of four standard deviations above the mean for the entire sequence. Second, the whitecap velocity and terminus were determined by tracking the centroid of the thresholded region. Last, a mean recovery time was computed for two regions measuring $0.5\text{ m} \times 0.5\text{ m}$ and located $\sim 0.5\text{ m}$ and 1.5 m behind the terminus and along the trajectory. The dashed line shows the T_n level of 0.3, which is used to define τ . This level of recovery was chosen to avoid uncertainty due to fluctuations sometimes observed after the initial recovery, such as the feature centred at $\sim 1.5\text{ s}$ in the top trace. (These features suggest secondary turbulence sources, such as bubbles².) b, Plot of τ versus C_b for 36 oceanic breaking waves during steady conditions (the wind speed at 10 m was $7\text{--}8\text{ m s}^{-1}$ and the net ambient heat flux, Q_{net} , determined from bulk methods was $120\text{--}130\text{ W m}^{-2}$ upwards). The vertical error bars show the average range of τ values, and the horizontal error bars reflect the average accuracy of the linear regression used to determine the component white-cap velocities. The correlation coefficient for the linear fit shown is 0.975.

specific heat of water, and δ is the thermal boundary layer thickness (taken as 0.5 mm for the observed wind conditions^{10,11}).

For an estimate of E , we use laboratory results showing that the rate of energy dissipation per unit crest length for quasi-steady¹⁷ and unsteady¹⁸ breaking waves is proportional to $\rho C^5/g$, where C is the breaking-wave phase speed and g is the gravitational acceleration. This relation has been employed in a number of recent efforts to model wave breaking in the open ocean^{15,18,19}. The area over which the dissipation occurs corresponds to the spatial extent of the turbulent wake. Most of the infrared wake signatures we have observed are roughly circular (Fig. 1), which suggests that the wake's dimension in the direction of propagation may be approximated by the length of the breaking crest. This assumption is consistent with previous measurements of the ratio of the breaking-crest length to the distance of wave advance during breaking². Therefore, an estimate of the energy dissipation rate per unit area is $E = 0.0032 (\rho C^5/g)/L_b$, where L_b is the breaking-crest length and the proportionality constant is that for unsteady spilling breakers¹⁸.

The variation of the skin-layer recovery rate with the energy dissipation rate is illustrated in Fig. 3 by plotting Q_r versus E . The restoring heat flux is correlated with the energy dissipation rate per unit area due to breaking. Note that Q_r is only part of the total breaking-induced heat flux as it does not include the contribution due to mixing below the skin layer. Furthermore, we estimate the effect of Q_r on the net ambient heat flux, Q_{net} , to be small because Q_r occurs only for the duration and over the extent of the breaking events, whereas Q_{net} is continuous in space and time. We emphasize that the utility of Q_r lies in its proportionality to the skin-layer recovery rate. Since Q_r is proportional to R_i , the results in Fig. 3 demonstrate that the skin-layer recovery rate in the wakes of breaking waves provides a measure of the energy dissipation rate due to breaking.

The field results are supported by complementary laboratory measurements of individual, mechanically generated breaking waves ranging in size from spilling to plunging. The technique we employed, in which the fraction of wave energy dissipated can be varied systematically, has been used extensively to investigate energy dissipation due to wave breaking^{20–23}. In the absence of wind, experiments using this technique have shown that turbulent kinetic energy remains measurable for 30–60 wave periods after breaking²². Figure 4a shows examples of three laboratory time series of T_n for breaking waves in the absence of wind. For these measurements, Q_r was computed using the R_i value given by the slope of the temperature deviation over 30 wave periods (here δ is taken as 3 mm for no wind²⁴). Figure 4b shows Q_r versus E for the

laboratory measurements. As in the field, Q_r and E are correlated. The difference in the magnitude of Q_r between these results, measured in a laboratory without wind, and the field results shown in Fig. 3 is a manifestation of the additional dependence of the skin-layer recovery process on ambient conditions¹³. The wind-generating facilities available during the laboratory experiments reported here permitted us to examine only low wind-speed conditions. Nonetheless, Q_r for wind speeds from 2 to 3 m s⁻¹ was ~ 300 W m⁻² (taking $\delta = 2$ mm for the given wind conditions^{10,11}), which is in the range of field values shown in Fig. 3.

These field and laboratory results show that infrared measurements provide a means of remotely quantifying the dynamics of surface-wave breaking. The rate of the skin-layer recovery in the wakes produced by breaking waves is correlated with the rate of energy dissipation due to wave breaking. Airborne applications of

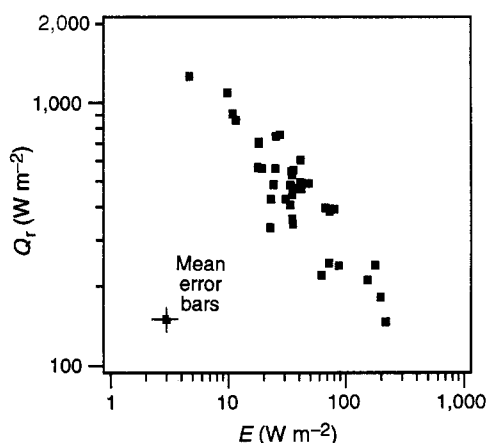


FIG. 3 The restoring heat flux, Q_r , versus the energy dissipation rate per unit area, E , for the oceanic breaking events shown in Fig. 2b. The correlation between Q_r and E implies that the skin-layer recovery rate, R_i , provides a measure of energy dissipation because Q_r is proportional to R_i . The error bars are analogous to those shown in Fig. 2b.

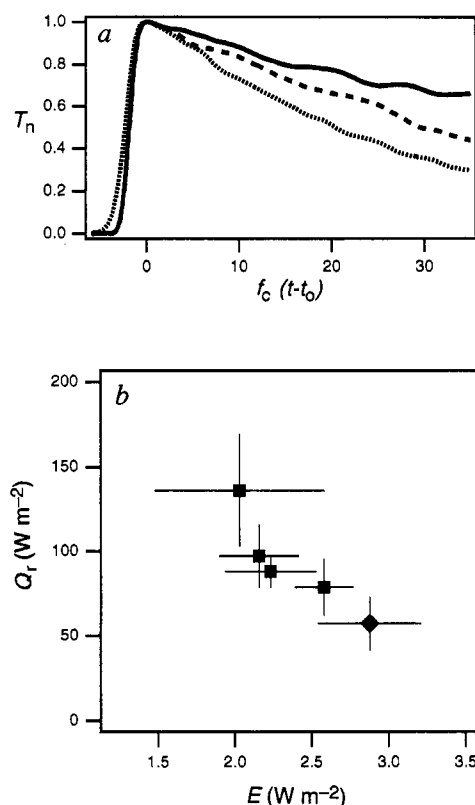


FIG. 4 a, Time series of T_n , the skin temperature deviation normalized by its maximum, for three laboratory breaking waves of different amplitude corresponding to a slope²² of 0.24 (dotted line), 0.25 (dashed line) and 0.27 (solid line). The abscissa is time expressed in wave periods since breaking, where t_c is the centre frequency of the wave packet, t is time, and t_0 is the time of breaking onset. b, The restoring heat flux, Q_r , versus the energy dissipation rate per unit area, E , for breaking waves generated mechanically in the laboratory for constant air–water temperature difference and no wind. Each point is an ensemble average of five repetitions and the error bars represent one standard deviation. The squares denote single breaking events and the diamond a double breaking event. The experiments were conducted in a 10-m long, 30-cm wide, and 40-cm deep wave tank located at the Canada Centre for Inland Waters. Measurements were made of the water and air temperatures and surface displacement, and infrared images were recorded using the same imager as in the field experiments. Breaking waves were produced by programming a hydraulic wave-maker to focus a dispersive wave packet at a fixed distance down the tank. The wave packet was synthesized from 32 sinusoidal components and had a centre frequency of 1.12 Hz, corresponding to a wavelength of 1.2 m. The quantity E is given by $0.5 \rho g C_g (\eta_0^2 - \eta_1^2) T_s / (L t_b)$, where C_g is the group velocity of the centre component of the packet, η_0^2 and η_1^2 are the surface displacement variances upstream and downstream of breaking, respectively, and T_s is the sampling period for the displacement variance²³. L is the along-tank length of the breaking region determined from the infrared images and t_b is the duration of the event as measured by a hydrophone.

this technique should provide remote measurements of energy dissipation due to wave breaking, which are necessary to improve wave prediction models used for operational sea-state forecasting. The capability to quantify the spatial and temporal characteristics of the skin-layer recovery has applications to other air-sea transfer processes that are directly affected by breaking-generated turbulence. Limitations in our understanding of ocean mixing and air-sea exchange mechanisms have hampered coupled models of the atmosphere and ocean used for global climate studies²⁵. The fluxes of heat and gas across the air-sea interface are enhanced by

the local increase in turbulence associated with wave breaking, and the skin-layer recovery rate has been shown to be directly related to properties of the free-surface turbulence. Furthermore, models of bubble-mediated gas transfer are critically dependent on accurate estimates of the surface area affected by wave breaking²⁶, which corresponds to the area of the detected wake. Although our results focus on remotely quantifying energy dissipation due to wave breaking, we conclude that infrared techniques may also lead to better predictions of air-sea heat and gas flux²⁷⁻²⁹. □

Received 5 December 1995; accepted 5 November 1996.

1. Banner, M. L. & Peregrine, D. H. *Annu. Rev. Fluid Mech.* **25**, 373-397 (1993).
2. Thorpe, S. A. *Prog. Oceanogr.* **35**, 315-352 (1995).
3. Melville, W. K. *Annu. Rev. Fluid Mech.* **28**, 279-321 (1996).
4. Komen, G. J. et al. *Dynamics and Modelling of Ocean Waves* (Cambridge Univ. Press, 1994).
5. Farmer, D. M. & Gemmrich, J. R. *J. Phys. Oceanogr.* **26**, 816-825 (1996).
6. Kitaigorodskii, S. A. *J. Phys. Oceanogr.* **14**, 960-972 (1984).
7. Katsaros, K. B. *Boundary-Layer Meteorol.* **18**, 107-127 (1980).
8. Robinson, I. S., Wells, N. C. & Charnock, H. *Int. J. Remote Sens.* **5**, 19-45 (1984).
9. McAlister, E. D. & McLeish, W. J. *Geophys. Res.* **74**, 3408-3414 (1969).
10. Wu, J. J. *J. Phys. Oceanogr.* **1**, 284-286 (1971).
11. Hill, R. H. *J. Phys. Oceanogr.* **2**, 190-198 (1972).
12. Ewing, G. & McAlister, E. D. *Science* **131**, 1374-1376 (1960).
13. Jessup, A. T., Zappa, C. J., Hesary, V., Loewen, M. R. & Skafel, M. G. in *Air-Water Gas Transfer* (ed. Jähne, B. & Monahan, E. C.) 601-610 (AEON Verlag & Studio, Hanau, 1995).
14. Eisner, L., Bell, E. E., Young, J. & Oetjen, R. A. *J. Opt. Soc. Am.* **52**, 201-209 (1962).
15. Phillips, O. M. *J. Fluid Mech.* **156**, 505-531 (1985).
16. Ding, L. & Farmer, D. M. *J. Phys. Oceanogr.* **24**, 1368-1387 (1994).
17. Duncan, J. H. *Proc. R. Soc. Lond. A* **377**, 331-348 (1981).
18. Melville, W. K. *J. Phys. Oceanogr.* **24**, 2041-2049 (1994).

19. Thorpe, S. A. *J. Phys. Oceanogr.* **23**, 2498-2502 (1993).
20. Chan, E. S. & Melville, W. K. *Proc. R. Soc. Lond. A* **417**, 95-131 (1988).
21. Melville, W. K. & Rapp, R. J. *Nature* **317**, 514-516 (1985).
22. Rapp, R. J. & Melville, W. K. *Phil. Trans. R. Soc. Lond. A* **331**, 735-800 (1990).
23. Loewen, M. R. & Melville, W. K. *J. Fluid Mech.* **224**, 601-623 (1991).
24. Katsaros, K. B., Liu, W. T., Businger, J. A. & Tillman, J. E. *J. Fluid Mech.* **83**, 311-335 (1977).
25. Barron, E. J. *Eos* **76**, 185-190 (1995).
26. Keeling, R. F. *J. Mar. Res.* **51**, 237-271 (1993).
27. Jähne, B., Libner, P., Fischer, R., Billen, T. & Plate, E. *J. Tellus* **41B**, 177-195 (1989).
28. Soloviev, A. V. & Schlüssel, P. *J. Phys. Oceanogr.* **24**, 1339-1346 (1994).
29. Haußecker, H. & Jähne, B. in *Air-Water Gas Transfer* (eds Jähne, B. & Monahan, E. C.) 775-784 (AEON Verlag & Studio, Hanau, 1995).

ACKNOWLEDGEMENTS. We thank the crew of RP FLIP and our colleagues at the University of Washington for their field assistance; M. G. Skafel, D. Doede (National Water Research Institute, Environment Canada) and M. A. O'Dor for their laboratory assistance; and F. Herr (ONR) for his early support and encouragement. This work was supported by the Applied Physics Laboratory, ONR, and NASA. M.R.L. was supported by a grant from the Natural Sciences and Engineering Research Council of Canada.

CORRESPONDENCE should be addressed to A.T.J. (e-mail: jessup@apl.washington.edu).

Emplacement of a large igneous province as a possible cause of banded iron formation 2.45 billion years ago

M. E. Barley*, A. L. Pickard* & P. J. Sylvester†

* Key Centre for Strategic Mineral Deposits, The University of Western Australia, Nedlands WA 6907, Australia

† Research School of Earth Sciences, The Australian National University, Canberra, ACT 0200, Australia

LATEST Archaean and earliest Palaeoproterozoic times (from 2.6 to 2.2 billion years ago) have generally been viewed as a largely quiescent period of Earth history; the geological record indicates the very slow deposition of pelagic and chemical sediments^{1,2}, and bears only a limited record of magmatic and tectonic activity³⁻⁵. Such quiescence is consistent with the contention that the Earth's main banded iron formations (BIFs)—finely laminated chemical sedimentary rocks, rich in iron oxide—formed slowly as oxygen abundances in the oceans gradually increased, thus reducing the capacity of sea water to retain dissolved iron⁶⁻¹⁰. Here we show that a large igneous province, comprising >30,000 km³ of dolerite, basalt and rhyolite, accompanied deposition of a Hamersley Province BIF 2.449 ± 3 million years ago. This observation indicates that Hamersley BIFs formed during a major tectono-magmatic event and were deposited very much faster than previously thought, at similar rates to (or faster than) modern pelagic sediments. Thus the largest Palaeoproterozoic BIFs, rather than simply reflecting a gradual increase in the oxygen content of the oceans during a period of tectonic quiescence, are more likely to have formed as a result of an increased supply of suboxic iron- and silica-rich sea water upwelling onto continental shelves during a pulse (or pulses) of increased submarine magmatic and hydrothermal activity.

Palaeoproterozoic BIFs are both a very significant indicator of Precambrian environmental conditions and the most important source of iron ores. The BIFs of the Hamersley Province in northwestern Australia, which form part of the 2.77-2.41-Gyr Mount Bruce Megasequence Set, are the world's thickest and most extensive (>50,000 km²) deposits of this type^{11,12}. They are finely laminated, chemical sedimentary rocks composed largely of microcrystalline quartz (chert) and iron oxide, with no modern counterparts. As a result, both their origin and tectonic context have been interpreted previously in nonuniformitarian terms as a conformable succession of rock units that formed continuously in a slowly subsiding intracratonic basin, continental shelf or submarine plateau, starved of clastic sediment and with little magmatic activity^{12,13}. The origin of iron and silica in such extensive finely laminated rocks is problematical and requires either that there was a direct magmatic-hydrothermal input to depositional basins¹⁴⁻¹⁷, or that continuously upwelling iron-rich waters from suboxic marine basins encountered more oxygen-rich waters on continental shelves^{6-8,17}. Iron oxides were simultaneously precipitated over large areas when a threshold oxygen activity was reached, perhaps as seasonal deposits in response to activity in the biosphere^{6,8,17,18}. The apparent absence of magmatic activity during deposition of the Hamersley BIFs is a key factor in the latter model and in arguments that Palaeoproterozoic BIFs reflect a gradual transition from suboxic to oxic conditions in the hydrosphere, rather than periods of enhanced magmatic and hydrothermal activity in marine basins⁶⁻¹⁰.

The oldest rocks in the Hamersley Province are flood basalts, rhyolites and fluvial, lacustrine and marine clastic sedimentary rocks of the 2.77-2.6-Gyr Chichester Range Megasequence (Fig. 1), which formed before, and during, the breakup and dispersal of a Late Archaean continent¹¹. These are overlain by the 2.6-2.41-Gyr Hamersley Range Megasequence. Initial deposition of BIFs, turbidites, organic and pyritic shales and carbonate sedimentary rocks (the Marra Mamba Supersequence Package) occurred on a continental shelf¹³. The Brockman Supersequence Package is dominated by BIFs with a cumulative thickness of >500 m with lesser fine tuff, turbidites, organic and pyritic shales, dolomite and

chert. Deposition occurred at $\sim 2,470$ Myr, on a clastic-starved open shelf or platform, possibly in a back-arc continental setting, with normal oceanic circulation¹¹. A tuff band in the Dales Gorge Member has a SHRIMP U–Pb in zircon age of $2,470 \pm 4$ Myr (ref. 19). (SHRIMP indicates Sensitive High Resolution Ion Microprobe.) The overlying Woongarra Supersequence comprises BIF, dolerite, basalt and rhyolite, also possibly deposited in a back-arc setting. These igneous rocks have traditionally been viewed as sills intruded after deposition and lithification of the BIFs^{12,20}. Following magmatism, BIFs and clastic sedimentary rocks of the Turee Creek Supersequence were deposited in a retroarc foreland basin¹¹.

The Woongarra Supersequence crops out over $\sim 40,000$ km². Its lower part (the Weeli Wolli Formation; Fig. 1) comprises 400 m of dolerite and basalt, interlayered with BIFs, shale and tuffaceous sedimentary rocks (up to 150 m). The igneous rocks outcrop poorly and have been interpreted as thick (up to 100 m) shallow-level sills. However, in an outcrop in Coondiner Gorge ($23^\circ 04' S$, $119^\circ 35' E$; Fig. 2), there is evidence for hyaloclastic flow-front breccias and BIFs deposited on top of a pillowed lava flow. This provides unequivocal evidence of eruption of basalt on to the sea floor during BIF deposition. Eruption on a submarine continental shelf is most likely because the mafic rocks are continental

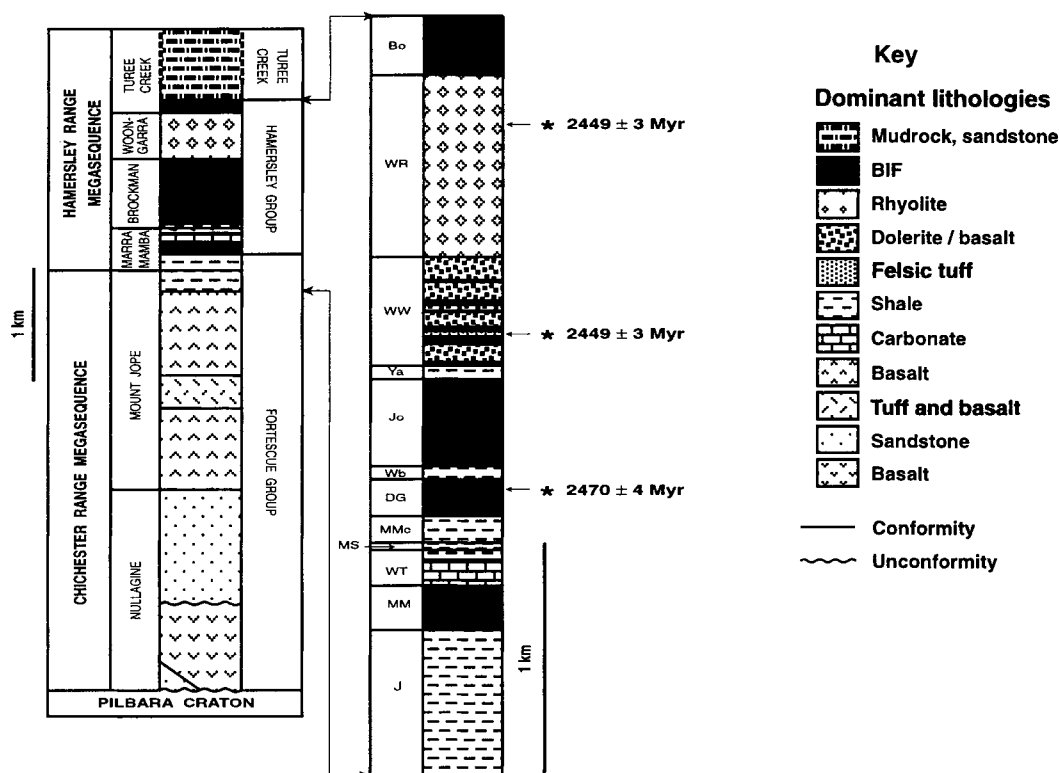


FIG. 1 Stratigraphic column of the Hamersley Province, and existing SHRIMP ages from the Hamersley Range Megasequence. DG, Dales Gorge; Jo, Joffre; WW, Weeli Wolli; WR, Woongarra Rhyolite; Bo, Boolgeeda; Ya, Yandi coogina; Wb, Whaleback; MMc, Mt McRae; MS, Mt Sylvia; WT, Wittenoom; MM, Marra Mamba; and J, Jeerinah. The dates with stars are used to define depositional rates for Hamersley BIF.

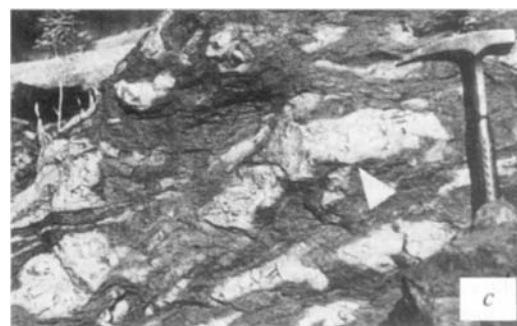
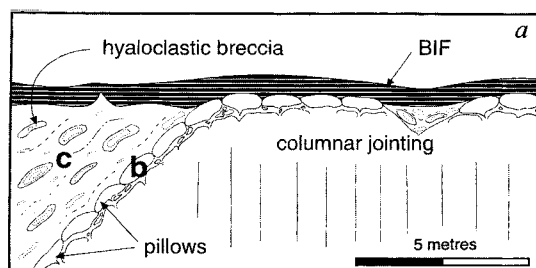


FIG. 2 a, Schematic diagram of basalt flow in the west wall of the Coondiner Gorge showing the geometry of lava flow with pillowed margin, hyaloclastic flow-front breccia and overlying BIF. b, Photograph showing details of pillows (arrowed) in flow front. c, Photograph showing pillows in the matrix of the hyaloclastic breccia (arrowed).

tholeiites, moderately enriched in all highly incompatible elements except Nb and Ta (Fig. 3a). The geochemistry of shales and tuffaceous sedimentary rocks indicates both mafic and felsic igneous provenances²¹. Fine-grained mafic tuffs were derived by sedimentary reworking of basaltic hyaloclastite, whereas tuffaceous sandstones and shales contain abundant quartz- and feldspar-phyric and vitric fragments and have trace-element abundances that indicate a rhyolitic source (Fig. 3a).

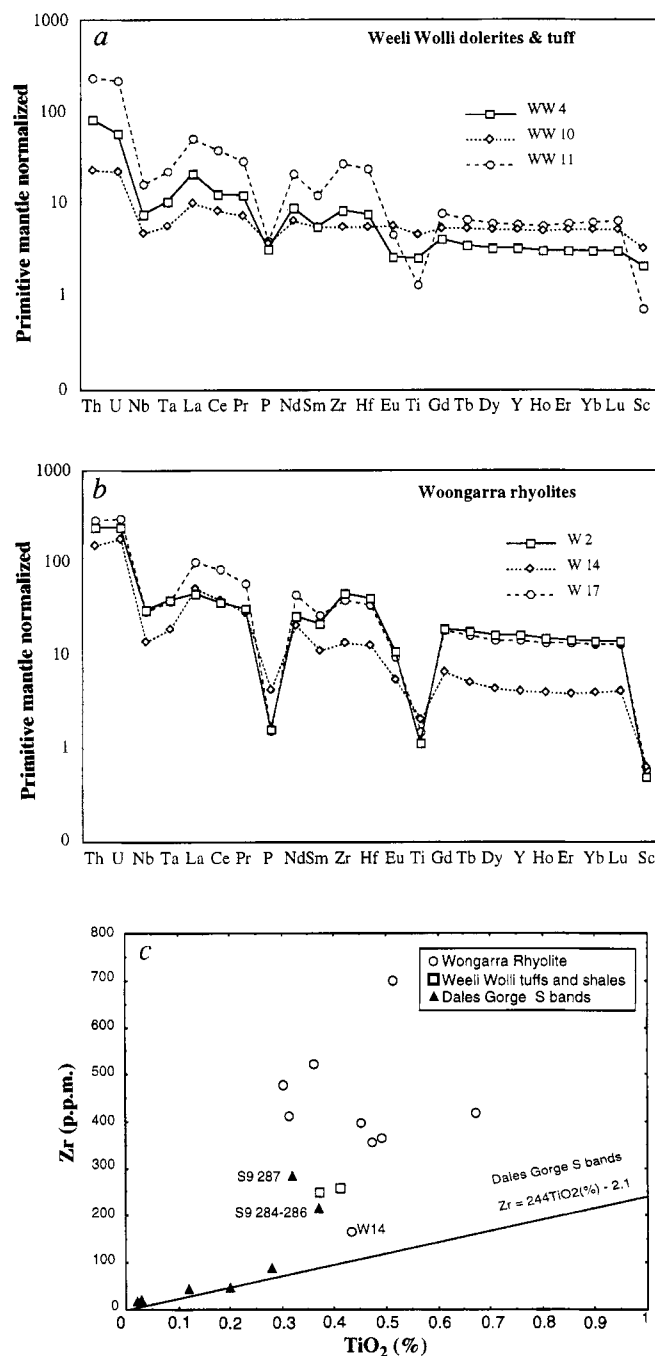


FIG. 3 Primitive mantle normalized trace-element diagram for: a, Weeli Wolli dolerite (WW4 and WW10) and felsic tuffaceous sandstone (WW11); b, Woongarra Rhyolite, (W2 and W17) and tuff (W14); c, Zr versus TiO_2 abundances for Woongarra Rhyolite, Weeli Wolli tuffaceous sandstones, and tuff and shale (S) bands from the Dales Gorge Member of the Brockman Supergroup Package (S 9 287 and S 9 284–286). Average S band data from Ewers and Morris³¹. Concentration factors and concentration data are available from P.J.S. and as Supplementary Information.

The upper part of the supersequence, the Woongarra Rhyolite, forms a 400-m-thick sheet of rhyolite. It comprises two composite cooling units separated by a thin unit of BIF, tuffaceous sedimentary rocks and, locally, dolerite. The rubby nature of outcrop and recrystallization make interpretation of the emplacement mechanism difficult. However, the common and widespread occurrence of cooling joints, amygdalae, flow banding, hyaloclastite, peperite, and tuffaceous turbiditic sedimentary rocks suggest shallow-water emplacement, either as lavas, or hot pyroclastic flows, or high-level sills into unconsolidated sediment. The lower rhyolite unit is almost aphyric with quartz paramorphs after tridymite indicating magmatic temperatures of $>900^\circ\text{C}$. The upper unit is quartz- and plagioclase-phyric with minor iron oxides and pseudomorphs after pyroxene. The rhyolite is similar to younger large-volume rhyolites with 70–74 wt% SiO_2 , 0.3–0.5% TiO_2 and relatively high total iron and low Al_2O_3 compared to average rhyolites²². They are also characterised by being relatively rich in high-field-strength elements such as Zr, with

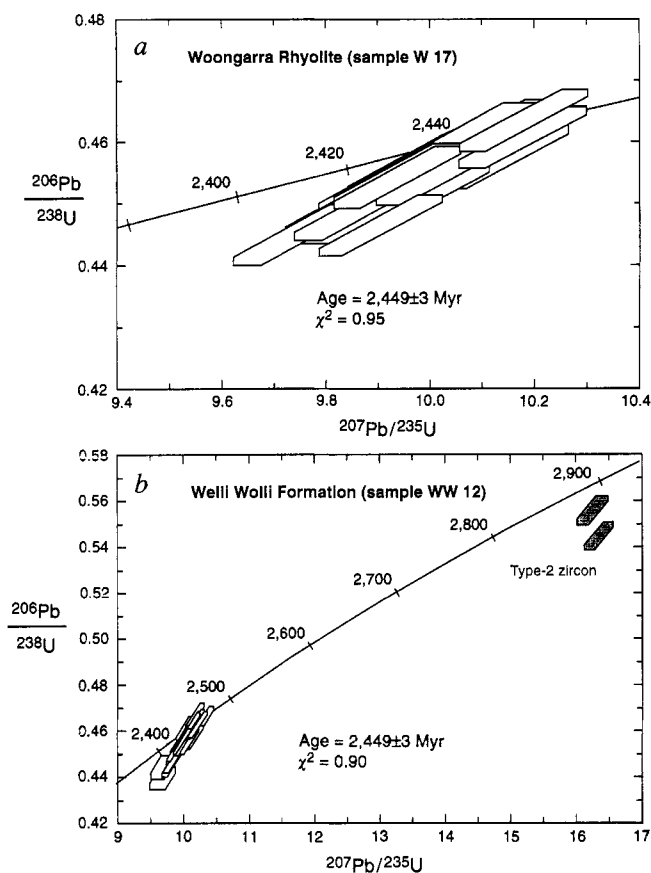


FIG. 4 Concordia diagrams showing the SHRIMP U–Pb systematics for zircons from a, Woongarra Rhyolite (sample W17), and b, Weeli Wolli Formation (sample WW12). Zircons extracted from these samples were analysed with the SHRIMP II mass spectrometer in Perth, using the CZ3 zircon standard (564 Myr; $^{206}Pb/^{238}U = 0.0914$) with precision (1σ) better than 1.04% for a and 1.2% for b. Initial Pb correction for the total Pb in each analysis was done by removing initial ^{206}Pb , ^{207}Pb and ^{208}Pb using the observed ^{204}Pb and assuming Pb of the Broken Hill isotopic composition. Errors are expressed at a 95% confidence level and ages in Myr are marked on the curve. In a, 24 spot analyses for 21 zircons plot as a single population giving an age of $2,449 \pm 3$ Myr for sample W17. In b, the largest subset of 28 spot analyses from 27 zircons (unshaded) with statistically indistinguishable $^{207}Pb/^{206}Pb$ define a magmatic age of $2,449 \pm 3$ Myr for sample WW12, older near-concordant analyses have light shading. Supplementary U–Pb analytical data are available from M.E.B. and as Supplementary Information.

moderate light-rare-earth element contents and negative Nb, Ta, Ti and Eu anomalies (Fig. 3b).

On the basis of volume ($>30,000 \text{ km}^3$), field relationships, volcanological characteristics, petrography and geochemistry, the Weeli Wolli dolerites and basalts and Woongarra Rhyolite represent a large igneous province which was emplaced during deposition of Weeli Wolli Formation BIFs. Emplacement of continental tholeiite magma preceded, and overlapped with, emplacement of high-temperature rhyolite, a common scenario in younger large igneous provinces associated with continental rifting and mantle plume activity²³. To constrain the timing of magmatism, a sample of the upper unit of the Woongarra Rhyolite, from Woongarra Gorge ($22^\circ 53' \text{ S}$, $117^\circ 06' \text{ E}$), and a felsic tuffaceous sandstone from the Weeli Wolli Formation, from the great Northern Highway west of Cathedral Gorge ($23^\circ 15' \text{ S}$, $119^\circ 34' \text{ S}$), have been dated by SHRIMP U–Pb in zircon geochronology^{24,25}. There are two types of zircon in the Woongarra Rhyolite sample, pale yellow-brown, clear, euhedral grains exhibiting faint euhedral zoning and three olive, slightly turbid, subhedral grains that lack visible internal zoning. All SHRIMP analyses plot as a single population with the same $^{207}\text{Pb}/^{206}\text{Pb}$ within experimental error (Fig. 4a), and the weighted mean age of $2,449 \pm 3 \text{ Myr}$ is interpreted as the magmatic age. The felsic tuffaceous sandstone sample (WW12) contains type-1 zircons which are pale-olive, clear, euhedral to subhedral grains, with some showing euhedral zoning, and a single darker, clear to slightly turbid, well rounded grain which is an older detrital grain. The dominant younger zircons are interpreted as juvenile volcanoclastic grains. All concordant to near-concordant grains from the younger population have the same $^{207}\text{Pb}/^{206}\text{Pb}$ within error (Fig. 4b) and their weighted mean $^{207}\text{Pb}/^{206}\text{Pb}$ age of $2,449 \pm 3 \text{ Myr}$ therefore represents a maximum depositional age for the sample.

Identical ages for detrital zircons in the tuffaceous sandstone and magmatic zircon in the Woongarra Rhyolite strongly support the interpretation that the Woongarra Supersequence represents a large igneous province. Furthermore, with a precision of $\pm 3 \text{ Myr}$, SHRIMP geochronology is not capable of resolving the depositional rate of the 500–100 m of BIF between the two sample points and, hence, this BIF must have been deposited considerably faster than the rate of $3\text{--}4 \text{ m Myr}^{-1}$ previously calculated for the Hamersley Range Megasequence¹ or the Griqualand West Sequence, South Africa². Also, more than 500 m of BIF and $\sim 100 \text{ m}$ of shale of the Brockman and Woongarra supersequences must have been deposited in the $\sim 20\text{-Myr}$ period between $2,470 \pm 4\text{-Myr}$ Brockman Supersequence tuff band¹⁹ and the $2,449 \pm 3\text{-Myr}$ Woongarra Rhyolite (Fig. 1). As BIFs are compacted and lithified sedimentary rocks, depositional rates for uncompacted Hamersley Province BIFs must have been faster than 30 m Myr^{-1} . This can be compared to the rate of 40 m Myr^{-1} estimated for modern uncompacted pelagic sediments²⁶. However, if sequence boundaries and organic shales represent periods of nondeposition and very slow deposition, respectively¹¹, then the Hamersley Province BIFs may have been deposited at even greater rates (between 100 and $1,000 \text{ m Myr}^{-1}$), as required by models which infer that paired chert-iron oxide laminae are seasonal deposits or varves^{8,17,18}.

Although the underlying Brockman Supersequence Package does not contain igneous rocks coeval with BIFs, some tuff bands in the Dales Gorge Member have trace-element signatures that are identical to the Woongarra Rhyolite tuffs and felsic tuffaceous sandstones in the Woongarra Supersequence (Fig. 3c). Tuff bands in the Joffre Member have geochemical characteristics of both mafic and felsic tuffs²⁷. These units have textures and sedimentary structures characteristic of turbidites and water-lain fine ash, indicating a major distal source of both juvenile basalt and rhyolite debris, similar to Woongarra Supersequence magmas, during deposition of Brockman Supersequence Package BIF. This suggests bimodal magmatism and associated hydrother-

mal activity within, or adjacent to, the marine basin flanking the Hamersley shelf at this time. Thus, there is mounting evidence that pulses of enhanced igneous and hydrothermal activity, related to a large igneous province (or provinces), may have accompanied both Brockman and Woongarra supersequence BIF deposition.

Evidence presented above supports arguments that the thickest and most extensive Palaeoproterozoic BIFs were coeval with pulses of crust and mantle-driven magmatism^{14–17} including emplacement of a large igneous province in the Woongarra Supersequence. Both Hamersley BIFs, and $\sim 2,470\text{--}2,430\text{-Myr}$ old BIFs in the southern African craton, which also contain extensive tuff bands, were being deposited when the Great Dyke was emplaced at $2,461 \pm 14 \text{ Myr}$ (refs 28, 29). The association of widespread ocean anoxia and ironstone accumulation with intense hydrothermal activity during periods of enhanced spreading at mid-ocean ridges and mantle plume activity is well documented for the Mesozoic³⁰. It can thus be argued that the massive flux of suboxic iron- and silica-rich sea water on to continental shelves during the Palaeoproterozoic may be a direct result of a period of enhanced submarine hydrothermal activity driven by a pulse (or pulses) of intense magmatic activity between 2,470 and 2,450 Myr ago, and that the Earth's largest BIFs formed as these waters mixed with the more oxygen-rich shallow hydrosphere. If this is the case, the evolution of the hydrosphere from suboxic conditions to more oxic conditions during the Palaeoproterozoic may not have been a gradual or steady state process, but instead may have been modulated by major tectono-magmatic events possibly preceeding the breakup of a Late Archaean supercontinent⁵. □

Received 9 August; accepted 30 October 1996.

1. Arndt, N. T., Nelson, D. R., Compston, W., Trendall, A. F. & Thorne, A. M. *Aust. J. Earth Sci.* **38**, 261–281 (1991).
2. Barton, E. S., Alterman, W., Williams, I. S. & Smith, C. B. *Geology* **22**, 343–346 (1994).
3. Condie, K. C. in *Archaean Crustal Evolution* (ed. Condie, K. C.) 85–120 (Elsevier, Amsterdam, 1994).
4. Lowe, D. R. in *The Proterozoic Biosphere; a Multidisciplinary Study* (eds Schopf, J. W. & Klein, C.) 67–76 (Cambridge Univ. Press, 1992).
5. Windley, B. F. *The Evolving Continents* (Wiley, Chichester, 1995).
6. Cloud, P. *Econ. Geol.* **68**, 1135–1143 (1973).
7. Holland, H. D. *The Chemical Evolution of the Atmosphere and Oceans* (Princeton Univ. Press, 1984).
8. Beukes, N. J. & Klein, C. in *The Proterozoic Biosphere; a Multidisciplinary Study* (eds Schopf, J. W. & Klein, C.) 146–151 (Cambridge Univ. Press, 1992).
9. Kasting, J. F. *Precamb. Res.* **34**, 205–229 (1987).
10. Des Marais, D. J. *Chem. Geol.* **114**, 303–314 (1994).
11. Blake, T. S. & Barley, M. E. *Tectonics* **11**, 1415–1425 (1992).
12. Trendall, A. F. & Blockley, J. G. *Bull. Geol. Surv. W. Aust.* **119** (1970).
13. Morris, R. C. & Horwitz, R. C. *Precamb. Res.* **21**, 273–297 (1983).
14. Isley, A. E. *J. Geol.* **103**, 169–185 (1995).
15. Ohmoto, H. *Geol. Soc. Aust. Abstr.* **37**, 324 (1994).
16. Tiara, A. et al. *Geol. Soc. Aust. Abstr.* **37**, 426 (1994).
17. Morris, R. C. *Precamb. Res.* **60**, 243–286 (1993).
18. Trendall, A. F. *Econ. Geol.* **68**, 1089–1097 (1973).
19. Trendall, A. F. et al. *3rd Int. Archaean Symp. Abstr.* 81–83 (Geoconferences, Perth, 1990).
20. Trendall, A. F. *Rep. Geol. Surv. W. Aust.* **42** (1995).
21. Davy, R. *Reed. Geol. Surv. W. Aust.* **1991/6** (1992).
22. Pankhurst, R. J. & Rapela, C. R. *Earth Planet. Sci. Lett.* **134**, 23–36 (1995).
23. Storey, M. *Nature* **377**, 301–308 (1995).
24. Compston, W., Williams, I. S. & Meyer, C. J. *Geophys. Res. (Suppl.)* **89**, B525–B534 (1984).
25. Williams, I. S. & Claesson, S. *Contrib. Mineral. Petrol.* **97**, 205–217 (1987).
26. Muller, P. J. & Mangini, A. *Earth Planet. Sci. Lett.* **51**, 94–114 (1980).
27. McConchie, D. M. thesis, Univ. Western Australia (1984).
28. Nisbet, E. G. *The Young Earth: an Introduction to Archaean Geology* (Allen & Unwin, London, 1987).
29. Hamilton, P. J. *J. Petrol.* **18**, 24–52 (1977).
30. Garzanti, E. *Geology* **21**, 105–108 (1993).
31. Ewers, W. E. & Morris, R. C. *Econ. Geol.* **76**, 1929–1953 (1981).

SUPPLEMENTARY INFORMATION is available on Nature's World-Wide Web site (<http://www.nature.com>) or as paper copy from Mary Sheehan at the London editorial office of Nature.

ACKNOWLEDGEMENTS. We thank R. Morris for providing samples of BIF shale and tuff bands, N. McNaughton and M. Dahl for geochronological assistance, and D. Groves, N. McNaughton and R. Morris for criticism. Zircon analyses were performed on a sensitive high-resolution ion microprobe mass spectrometer (SHRIMP 11) operated by a consortium consisting of Curtin University of Technology, the Geological Survey of Western Australia and the University of Western Australia with support from the Australian Research Council. Trace-element analyses were performed by solution nebulization-ICP-MS at the Research School of Earth Sciences, Australian National University. M.E.B. also acknowledges a visiting fellowship at the Research School of Earth Science.

CORRESPONDENCE should be addressed to M.E.B. (e-mail: mbarley@geol.uwa.edu.au).

Conifer root discrimination against soil nitrate and the ecology of forest succession

Herbert J. Kronzucker, M. Yaeesh Siddiqi
& Anthony D. M. Glass

Department of Botany, University of British Columbia, Vancouver,
British Columbia V6T 1Z4, Canada

THE high incidence of failure when late-successional conifer species are replanted on disturbed forest sites is a considerable problem¹⁻³. Here we advance a hypothesis that might explain many of these reforestation problems on a physiological basis, within the framework of forest succession. It is known that the chemical speciation of inorganic nitrogen in forest soils changes from predominantly ammonium (NH_4^+) in late-successional (mature forest) soils to mostly nitrate (NO_3^-) after disturbances such as clearcut harvesting²⁻⁶. The capacity of plant roots to take up and use these two sources of nitrogen is therefore very important for species establishment on successional different sites. We have used kinetic and compartmental-analysis techniques with the radiotracer ^{15}N to compare the efficiency of nitrogen acquisition from NH_4^+ and NO_3^- sources in seedlings of white spruce, an important late-successional conifer. We found that uptake of NH_4^+ was up to 20 times greater than that of NO_3^- from equimolar solution, cytoplasmic concentration of NH_4^+ was up to 10 times greater than that of NO_3^- , and physiological processing of NO_3^- was much less than that of NH_4^+ . This reduced capacity to use NO_3^- is thought to present a critical impediment to seedling establishment on disturbed sites, where species better adapted to NO_3^- would have a significant competitive advantage.

Inorganic nitrogen is available to plants in soil solution either as NO_3^- or as NH_4^+ . Physiological competition at the root level for this resource may have profound effects on relative species performance in the field⁷. Most agricultural species and species confined to poor-quality ruderal or pioneer soils can use either nitrogen source⁸. The ratio of uptake of NO_3^- to NH_4^+ in these species is typically close to 1 (refs 9, 10). Nevertheless, growth on NO_3^- is frequently superior to growth on NH_4^+ (ref. 11), and, on NH_4^+ , toxicity is sometimes observed¹². For this reason, such

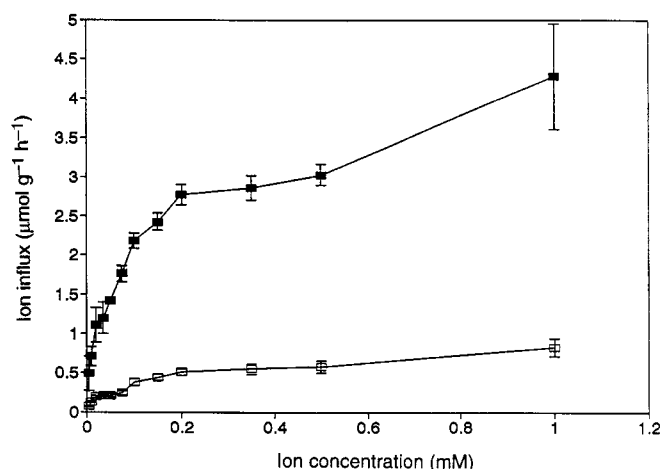


FIG. 1 Comparative concentration dependence of NO_3^- influx (open symbols) and NH_4^+ influx (filled symbols) in roots of intact white-spruce seedlings in the range of environmentally relevant nitrogen concentrations⁶. All plants were pretreated at $100\ \mu\text{M}$ $[\text{NO}_3^-]$ for 3 days. Error bars indicate s.e. ($N \geq 12$).

species are sometimes said to be 'nitrophilous'². Another category of species, including rice¹³, ericaceous species and many conifers^{2,12}, occur naturally on soils enriched in NH_4^+ and organic nitrogen^{2-6,13}, and appear not to suffer ammonium toxicity^{2,12-15}. Nitrate concentrations in these soils are virtually undetectable as a result of inhibited nitrification^{2,4,5,13} and/or preferential microbial acquisition of NO_3^- (ref. 31). From an evolutionary perspective, such species might be expected to perform more efficiently on reduced nitrogen than the aforementioned nitrophiles, and to deal successfully with relatively high external NH_4^+ . By contrast, the scarcity of NO_3^- would render efficient NO_3^- assimilation virtually irrelevant. Indeed, conifers are reported to grow much better on NH_4^+ than NO_3^- (refs 2, 16, 17).

In most forest soils, available nitrogen changes markedly with the stage of successional development³. After disturbance, soil pH generally rises and a new microbial environment appears, which converts soil nitrogen from predominantly NH_4^+ to mostly NO_3^- (refs 2-5). Late-successional conifers therefore become poor competitors for inorganic nitrogen, and the site becomes dominated by nitrophiles^{2,3,18-20}. Factors other than adaptation to nitrogen source, such as sun intolerance¹⁸ and inherently slow growth in conifers²¹, are also thought to influence this successional pattern. In boreal, montane and subalpine environments, early-successional species, such as aspen, frequently invade sites formerly dominated by spruce after clearcut harvesting, despite replanting with spruce seedlings. Reforestation failure of this sort is substantial in many parts of North America; in British Columbia alone, more than 1.5 million ha of productive forest lands have been classified as failed replantings¹. Thus, both economically and ecologically, the identification of a possible physiological determinant of reforestation success is of considerable interest.

We have used the radiotracer ^{15}N to assess the differential capacities for the utilization of NO_3^- and NH_4^+ in seedlings of white spruce, a late-successional conifer. Rates of nitrogen uptake were up to 20 times higher for NH_4^+ than for NO_3^- (refs 22, 23). Although V_{max} values under those conditions were vastly different ($\approx 0.1\ \mu\text{mol g}^{-1}\text{ h}^{-1}$ versus $\approx 2\ \mu\text{mol g}^{-1}\text{ h}^{-1}$, respectively), K_m values were very similar ($15\text{--}20\ \mu\text{M}$), possibly indicating differences in the numbers of the respective transporter proteins in the

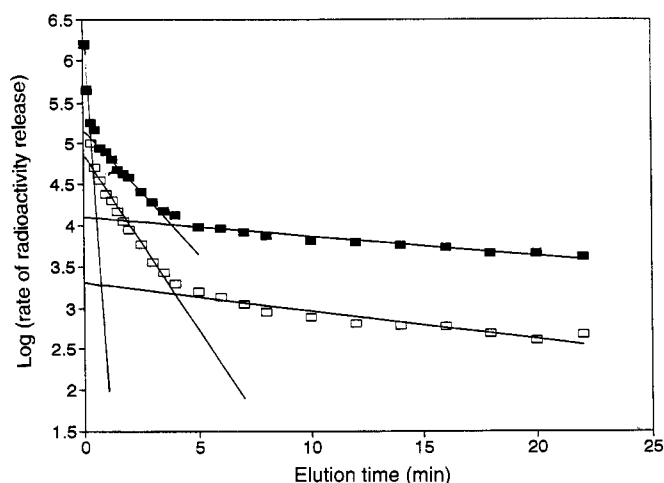


FIG. 2 Combined semi-logarithmic plots of the rate of release of $^{15}\text{NO}_3^-$ (open symbols) and $^{15}\text{NH}_4^+$ (filled symbols) in $[\log (\text{c.p.m. released})\ \text{g}^{-1}\text{ min}^{-1}]$ versus time of elution for roots of intact white-spruce seedlings at $100\ \mu\text{M}$ NO_3^- or NH_4^+ . Plots include linear regression lines for the three logarithmic phases of efflux for both sets of experiments. Counts eluting from root tissues were corrected for differences in specific activity of the radiotracer to allow for direct comparison of y-intercepts.

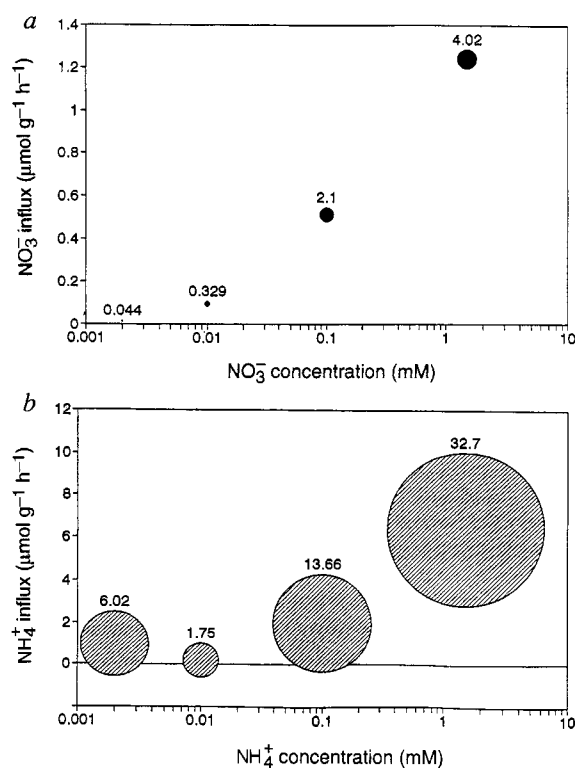


FIG. 3 Comparative plots of influx of a, NO_3^- and b, NH_4^+ and cytoplasmic pool size at four different external concentrations of the two nitrogen sources. Sizes of shaded circles are calculated to be proportional to one another and represent relative cytoplasmic concentrations (absolute values are indicated (in mM) above). Centres of the circles (y-axis values) indicate rates of unidirectional influx of NO_3^- and NH_4^+ under the given conditions as determined by efflux analysis. NO_3^- or NH_4^+ were added 3 days before (and during) the efflux experiment^{24,27} (at 10 μM , 100 μM or 1.5 mM). In the case of nitrogen-deprived plants, NO_3^- and NH_4^+ were withheld from growth and pretreatment solutions, and only added (at 10 μM) during labelling and elution to make possible the monitoring of fluxes and the estimations of compartmental concentrations.

plasma membrane, rather than differences in the affinities of the transporters for the two substrates.

As in most species, NO_3^- uptake in spruce is enhanced ('induced') by previous exposure to NO_3^- (ref. 24). However, maximal induction of NO_3^- transport required three days of NO_3^- exposure (at 100 μM NO_3^-), but only several hours in cereals²⁴. Indeed aspen, a typical pioneer tree species^{12,21}, does not require any prior exposure to NO_3^- to achieve maximal uptake (data not shown). This represents further evidence for poor adaptation to utilization of NO_3^- in white spruce. A direct comparison of NO_3^- and NH_4^+ influx in NO_3^- -induced seedlings still shows four- to fivefold higher influx for NH_4^+ than for NO_3^- (Fig. 1). A preference of a similar magnitude for NH_4^+ over NO_3^- has been documented in longer-term net-flux estimations in several conifer species^{2,14,15,25,26}. Furthermore, spruce also demonstrates low rates of NO_3^- reduction and subsequent metabolic processing²⁴.

At the subcellular level, efflux analysis was used to determine kinetic constants for the exchange of NO_3^- and NH_4^+ , to quantify the extent of efflux under varying nitrogen supply, as well as to estimate the cytoplasmic concentrations of the two ions. Three kinetically distinct subcellular compartments were revealed for both nitrogen sources (Fig. 2): (1) a root-surface film; (2) an adsorptive component of the cell wall; and (3) the root-cell cytoplasm^{24,27,28}. Half-lives of exchange for these compartments for NO_3^- were 2 s, 20 s and 7 min, respectively, and for NH_4^+ they

were 2 s, 30 s and 14 min, respectively^{24,27,28}. Efflux analysis confirmed the strong preference for NH_4^+ . Plots of $^{13}\text{NO}_3^-$ and $^{13}\text{NH}_4^+$ efflux clearly show significantly lower y-intercepts for both the cell-wall and the cytoplasmic regression lines for NO_3^- compared to NH_4^+ (Fig. 2). The differences indicate smaller fluxes and a smaller accumulation capacity in the cell-wall free space and the cytoplasm of spruce roots for NO_3^- compared to NH_4^+ (refs 24, 27). We measured the changes in influx and cytoplasmic concentrations of NO_3^- was consistently five- to eightfold lower than that of NH_4^+ . In seedlings previously starved of nitrogen, NO_3^- accumulation in the cytoplasm was 140-fold lower, an effect which can be attributed to the fact that NO_3^- uptake requires induction, whereas NH_4^+ uptake does not. Cytoplasmic NH_4^+ in spruce roots was similar in magnitude to levels in rice, which is also adapted to NH_4^+ soils¹³, and to NO_3^- levels in agricultural species^{24,27}, which have superior growth on NO_3^- (refs 11, 12, 18–20).

Mycorrhization of the root system, as would be expected for spruce seedlings in the field, has been shown to enhance NH_4^+ absorption rates in conifers, but does not significantly affect rates of NO_3^- uptake^{29,30}. The preference for NH_4^+ , as observed in our non-mycorrhizal plant material, should therefore be accentuated even further in the field. In our opinion, given the pronounced inherent differences in the physiological utilization capacities for NH_4^+ and NO_3^- as sources of nitrogen, it is not surprising that reforestation problems are encountered with species such as white spruce on disturbed sites, where NO_3^- is the predominant nitrogen source and NH_4^+ is in short supply. □

Methods

Influx experiments. Three-month-old nursery-grown seedlings of *Picea glauca* (Moench) Voss²⁷ were transferred to hydroponic tanks containing aerated 1/10-strength modified nitrogen-free Johnson's solution. Growth and pretreatment conditions were as described²⁷. NO_3^- (as $\text{Ca}(\text{NO}_3)_2$) or NH_4^+ (as $(\text{NH}_4)_2\text{SO}_4$) were added during uptake at the indicated concentrations. The radiotracer ^{13}N was provided by the Tri-University Meson Facility (TRIUMF) at the University of British Columbia, and $^{13}\text{NH}_4^+$ and $^{13}\text{NO}_3^-$ were generated^{13,27}. Roots of intact seedlings were equilibrated for 5 min in non-labelled solutions chemically identical to the uptake solutions, transferred to uptake vessels containing $^{13}\text{NH}_4^+$ or $^{13}\text{NO}_3^-$ labelled solution for 10 min, and postwashed in non-labelled solution for 2 min (for NO_3^-) or 3 min (for NH_4^+) to desorb tracer contained in the free space^{23,25}. Seedling roots were then excised from shoots, the roots were spun in a low-speed centrifuge for 30 s to remove surface liquid, and the fresh weights of roots and shoots were determined. The radioactivities of roots and shoots were determined in a Packard γ -counter (Minaxi 8, Auto- γ 5000 Series). Using the value for specific activity ($^{13}\text{N}/(^{13}\text{N} + ^{14}\text{N})$) of the loading solution and the total fresh root weight of each seedling, NO_3^- or NH_4^+ fluxes were calculated and expressed in $\mu\text{mol g}^{-1} \text{h}^{-1}$ (refs 22, 23). Experiments were repeated four times. Each experimental treatment consisted of three seedling samples (minimum root mass was 3 g fresh weight per sample). Data from several experiments were pooled ($N \geq 12$) for calculations of means and standard errors. These values were used for plotting the representative concentration-dependent curves, as well as for calculating V_{max} and K_m values as described²².

Compartmental analysis. Roots of intact seedlings (grown at steady-state provision of 10 μM , 100 μM or 1.5 mM NO_3^- or NH_4^+) were equilibrated in non-labelled preloading solution for 5 min before transfer to the ^{13}N -loading solution. Roots were then immersed in ^{13}N -labelled loading solution for 35 min (for NO_3^-) to 60 min (for NH_4^+) to bring the cytoplasmic phase to a specific ^{13}N activity close to that of the loading solution^{27,28}. The seedlings were transferred to efflux funnels¹³, and the roots eluted successively with small portions of non-labelled solution for varying times. With $t = 0$ as the time of transfer from loading to washing solution, and $t_{\text{final}} = 22$ min for the final elution, the time periods for the 25 successive washes were: 5 s (2 $\times$), 10 s (2 $\times$), 15 s (6 $\times$), 30 s (4 $\times$), 1 min (4 $\times$), 2 min (7 $\times$). The eluates were then counted in a Packard γ -counter as above. Roots were excised from the shoots after the final elution, were spun for 30 s, and plant organs were weighed and counted. Treatment of data was as described^{24,27}. Experiments were repeated 4–7 times, with two replicates each. Standard errors for various derived parameters (half-lives, fluxes, pool sizes) were within 15% of the means ($N \geq 8$). Representative experiments were chosen for semi-logarithmic plots of the rate of ^{13}N release versus elution time (Fig. 2). Because the specific activity in the plant compartments during elution is declining exponentially, the logarithm of the rate of release of radioactivity from the plant tissue can be plotted against elution time^{13,4,27,28}. Linear regression on the semi-logarithmic plots was then used to resolve separate phases. The slopes of the regression lines, after conversion to natural logarithm, yielded kinetic exchange constants (k) for the respective phases, which could be expressed as

half-lives of exchange ($t_{1/2} = 0.693/k$). The intercept with the ordinate of the regression line for the presumed cytoplasmic phase (that is, the rate of ^{15}N release from the slowest-exchanging phase at time zero) indicates the size of the cytoplasmic NO_3^- or NH_4^+ pool^{13,28}. Cytoplasmic concentrations of NO_3^- or NH_4^+ were calculated from the quotient of the integrated rate of ^{15}N release during five times the half-life of cytoplasmic exchange and the ratio of efflux to all fluxes removing $^{15}\text{NO}_3^-$ or $^{15}\text{NH}_4^+$ from the cytoplasm, and assuming 5% for cell volume occupied by the cytoplasm^{24,27,28}.

Received 26 April; accepted 28 October 1996.

1. Ministry of Forests, British Columbia, Canada. Annual report 1991–1992.
2. Lavoie, N., Vézina, L.-P. & Margolis, H. *Tree Physiol.* **11**, 171–183 (1992).
3. Jobidon, R., Thibault, L.-P., Fortin, J. A. *For. Ecol. Mgmt* **29**, 295–310 (1989).
4. Likens, G. E., Borman, F. H. & Johnson, N. M. *Science* **163**, 1205–1206 (1969).
5. Vitousek, P. M., Matson, P. A. & Van Cleve, K. *Plant Soil* **115**, 229–239 (1989).
6. Vogt, M. & Edmonds, R. L. *Northwest Sci.* **56**, 83–89 (1982).
7. Martin, M. P. D. & Snaydon, R. W. J. *J. Appl. Ecol.* **19**, 263–272 (1982).
8. Vessey, J. K., Henry, L. T., Chailouf, S. & Raper, C. D. Jr. *J. Plant. Nutr.* **13**, 95–116 (1990).
9. Lee, R. B., Purves, J. V., Ratcliffe, R. G. & Saker, L. R. J. *Exp. Bot.* **43**, 1385–1396 (1992).
10. Fentem, P. A., Lea, P. L. & Stewart, G. R. *Plant Physiol.* **71**, 502–506 (1983).
11. Malhi, S. S., Nyborg, M., Jahn, H. G. & Penny, D. C. *Plant Soil* **105**, 231–239 (1988).
12. Pearson, J. & Stewart, G. R. *New Phytol.* **125**, 283–305 (1993).
13. Wang, M. Y., Siddiqi, M. Y., Ruth, T. J. & Glass, A. D. M. *Plant Physiol.* **103**, 1249–1258 (1993).
14. Scheromm, P. & Plassard, C. *Plant Physiol. Biochem.* **26**, 261–269 (1988).
15. Marschner, H., Häussling, M. & George, E. *Trees* **5**, 14–21 (1991).
16. McFee, W. W. & Stone, E. L. *Jr Soil. Sci. Soc. Am. Proc.* **32**, 879–884 (1968).
17. van den Driessche, R. *Plant Soil* **34**, 421–439 (1971).
18. Smirnov, N., Todd, P. & Stewart, G. R. *Ann. Bot.* **54**, 363–374 (1984).
19. Stewart, G. R., Mann, A. F. & Fentem, P. A. *Physiol. Plant.* **74**, 26–33 (1988).
20. Pearson, J., Clough, E. C. M. & Kershaw, J. L. *Ann. Sci. For.* **46** (suppl.), 663–665 (1989).
21. Chapin, F. S., van Cleve, K. & Tyron, P. R. *Oecologia* **69**, 238–242 (1986).
22. Kronzucker, H. J., Siddiqi, M. Y. & Glass, A. D. M. *Plant Physiol.* **109**, 319–326 (1995).
23. Kronzucker, H. J., Siddiqi, M. Y. & Glass, A. D. M. *Plant Physiol.* **110**, 773–779 (1996).
24. Kronzucker, H. J., Glass, A. D. M. & Siddiqi, M. Y. *Planta* **196**, 683–690 (1995).
25. Flaig, H. & Mohr, H. *Physiol. Plant.* **84**, 568–576 (1992).
26. Knoepp, J. D., Turner, D. F. & Tingey, D. T. *For. Ecol. Mgmt* **59**, 179–191 (1993).
27. Kronzucker, H. J., Siddiqi, M. Y. & Glass, A. D. M. *Planta* **196**, 674–682 (1995).
28. Kronzucker, H. J., Siddiqi, M. Y. & Glass, A. D. M. *Plant Physiol.* **109**, 481–490 (1995).
29. Alexander, I. J. in *Nitrogen as an Ecological Factor* (eds Lee, J. A., McNeill, S. & Ronson, I. H.) 69–93 (Blackwell, Oxford, 1983).
30. Plassard, C., Barry, D., Eltrop, L. & Mousin, D. *Can. J. Bot.* **72**, 189–197 (1994).
31. Stark, J. M. & Hart, S. C. *Nature* **385**, 61–64 (1997).

ACKNOWLEDGEMENTS. We thank P. Culbert, R. D. Guy, U. Heber, T. J. Ruth, B. Touraine, J. J. Vidmar and J. G. Worrall for technical help and discussion.

CORRESPONDENCE and requests for materials should be addressed to H.J.K. (e-mail: herbertk@botany.ubc.ca).

High rates of nitrification and nitrate turnover in undisturbed coniferous forests

John M. Stark* & Stephen C. Hart†

* Department of Biology and the Ecology Center, Utah State University, Logan, Utah 84322-5305, USA

† School of Forestry, College of Ecosystem Science and Management, Northern Arizona University, Flagstaff, Arizona 86011-5018, USA

THE importance of nitrate (NO_3^-) in the internal nitrogen cycle of undisturbed coniferous ecosystems has not been widely recognized^{1,2}. Nitrate concentrations in soils from these forests tend to be low, and assays measuring net nitrification usually show exceedingly slow rates^{3,4}. It may be, however, that microbial assimilation of NO_3^- is substantial in these soils, and that net nitrification rates greatly underestimate gross rates⁵. Here we use a ^{15}N isotope-dilution technique in intact soil cores to measure gross rates of nitrification and microbial assimilation of NO_3^- in eleven undisturbed forest ecosystems of New Mexico and Oregon. We found that gross nitrification rates were surprisingly high in all of the forests examined. Net nitrification rates poorly predicted gross rates because the soil microbial communities had the capacity to assimilate almost all of the NO_3^- produced. To our

knowledge, this is the first report of gross nitrification and NO_3^- assimilation rates in intact soil samples from a large number of contrasting forest ecosystems. Our results contradict previous assumptions that nitrification rates are low in mature coniferous forests and suggest that current models greatly underestimate the role of the microbial community in preventing NO_3^- loss.

Studies examining nitrogen retention in forest ecosystems have focused on net nitrification, net mineralization, microbial assimilation of ammonium (NH_4^+), and plant uptake; however, microbial assimilation of NO_3^- has been largely ignored. Early studies showed that soil microbial communities prefer NH_4^+ to NO_3^- as a source of nitrogen⁶, and measurable quantities of NH_4^+ are almost always present in soils. Therefore, it has generally been assumed that NH_4^+ will be the nitrogen source for microbes, and that microbial assimilation of NO_3^- will be minimal^{7,8}.

Microbial assimilation of NO_3^- has been discounted as a significant process controlling NO_3^- pool sizes and as a mechanism of nitrogen retention following disturbance in forest ecosystems^{3,9,10}, but we considered that substantial amounts of microbial assimilation of NO_3^- might occur in these systems because: high rates of carbon addition to soils are likely to result in nitrogen limitation to the microbial biomass; high spatial variability in carbon inputs is likely to result in microsites of mineralization and nitrification adjacent to microsites of intense immobilization; and fungal populations dominating these soils are capable of translocating NO_3^- from microsites with high mineralization and nitrification to microsites with high carbon availability.

We evaluated the importance of nitrification and microbial assimilation of NO_3^- by measuring rates in eleven forest soils along an elevational transect in the Tesuque watersheds of northern New Mexico^{3,9,10} and a latitudinal transect (at about 44°N) extending 220 km from the Oregon coast to the east side of the Cascade Mountains in central Oregon^{11,12} (Table 1). The ecosystems along these two transects represent a wide range of forest ecosystems and span almost the entire range of above-ground net primary production that occurs in forests of North America (1 to $13\text{ Mg ha}^{-1}\text{ yr}^{-1}$)^{11,12}.

Gross rates of nitrification and NO_3^- consumption were measured in late spring (May and June) and late summer (August) using $^{15}\text{NO}_3^-$ isotope dilution¹³. Rates were measured during *in situ* incubation of intact core samples and homogenized samples from the 0–15-cm mineral soil layer. Homogenized samples were incubated with and without acetylene (at 10 kPa): to verify the validity of isotope dilution measurements; to evaluate the relative importance of autotrophic compared with heterotrophic nitrification; and to determine how much NO_3^- consumption was due to denitrification. Rates of microbial assimilation of NO_3^- were also verified by measuring microbial biomass ^{15}N and total soil organic ^{15}N at the end of the incubations.

Gross nitrification rates in intact soil cores were high at all eleven forest sites (Table 2), ranging from $25\text{ mg N m}^{-2}\text{ d}^{-1}$ in the New Mexico ponderosa pine site during summer, to $>300\text{ mg N m}^{-2}\text{ d}^{-1}$ in the Douglas-fir site during spring. These rates are one to two orders of magnitude higher than rates of nitrogen input from litterfall^{3,11}, indicating that cycling of NO_3^- through the soil microbial community is extremely rapid relative to plant nitrogen uptake. High nitrification rates occur in these forests in spite of low soil pH (Table 1), low nitrogen deposition rates, and low availability of nitrogen. Wet deposition of nitrogen at each of these sites averaged $<2\text{ kg N ha}^{-1}\text{ yr}^{-1}$ during the past decade¹⁴.

Comparison of nitrification rates determined by isotope dilution in homogenized samples with rates determined by acetylene inhibition verified that rate estimates from isotope dilution are reliable ($k_a = -0.08 + 0.97k_N$, where k_N is the gross nitrification rate determined by ^{15}N isotope dilution (in $\text{mg N kg}^{-1}\text{ d}^{-1}$), and k_a is the rate calculated from the difference between net NO_3^- consumption with and without acetylene; $r^2 = 0.67$; $n = 34$). Acetylene inhibition also showed that, in spite of low soil pH, almost all of the nitrification was due to the activity of autotrophic

rather than heterotrophic nitrifier populations. Only soil at the Oregon ponderosa pine site had significant gross nitrification rates following exposure to acetylene ($0.30 \text{ mg N kg}^{-1} \text{ d}^{-1}$; $P < 0.025$).

In spite of high nitrification rates, NO_3^- concentrations were low in almost all of the forest soils examined (Table 1). With the exception of the pinyon/juniper and red alder/Douglas-fir sites, ambient soil NO_3^- concentrations were $< 0.6 \text{ mg N kg}^{-1}$; NO_3^- concentrations in more than half of the plots at the Oregon ponderosa pine and western hemlock/sitka spruce sites were below detection limits ($< 0.05 \text{ mg N kg}^{-1}$). Because of small NO_3^- pools and high flux rates, mean residence times for NO_3^- pools ranged from $< 1 \text{ h}$ in the Oregon ponderosa pine site in spring and the western hemlock/sitka spruce site in summer, to 1.47 days in the western juniper site in summer (Table 2).

Net nitrification rates provided little indication of the high rates of nitrification that were occurring at the sites. Mean net nitrification rates were negative in 13 of 19 cases (Table 2). These net rates, determined during 1-day incubations, encompass the range of net nitrification rates determined over longer incubation periods in other undisturbed mature temperate forests (-2 to $41 \text{ mg N m}^{-2} \text{ d}^{-1}$) (calculated from data in ref. 15, assuming a 180-day active period). Net nitrification poorly predicted gross nitrification rates because in most soils all of the NO_3^- produced was rapidly assimilated by microorganisms.

Measurements made of extractable microbial ^{15}N , soil organic ^{15}N , and denitrification verified that the NO_3^- consumption in these soils was almost completely due to microbial assimilation. Isotope dilution estimates of NO_3^- consumption were strongly correlated with rates of accumulation of ^{15}N in the chloroform-labile microbial biomass nitrogen pool and in the soil organic nitrogen fraction ($r^2 > 0.65$; slopes of both regression lines were not significantly different from 1.0 at $P \leq 0.05$). Denitrification,

measured in homogenized samples exposed to C_2H_2 , accounted for $< 4\%$ of NO_3^- consumption at all sites except the Douglas-fir site in spring, where it accounted for 17%.

In intact core samples, NO_3^- consumption was tightly coupled to gross nitrification rates (Fig. 1), suggesting that NO_3^- assimilation was limited by NO_3^- availability. At low and moderate nitrification rates, NO_3^- consumption exceeded nitrification (as indicated by points above the 1:1 line). Addition of $^{15}\text{NO}_3^-$ may have stimulated consumption rates in these samples¹³, and thus consumption rates in unamended soils may be lower than the rates shown here. Nevertheless, our results demonstrate that the microbial communities in most of the ecosystems have the capacity to assimilate even more NO_3^- than is produced by nitrification.

Rates of NO_3^- consumption were consistently lower than nitrification only when nitrification rates were very high (Fig. 1). These data apparently represent soils with high NH_4^+ availability relative to carbon availability, which stimulates nitrification rates and suppresses NO_3^- assimilation. For example, soils tended to have higher rates of NO_3^- production than consumption at the red alder/Douglas-fir site, where nitrogen fixation increased the availability of NH_4^+ , and at the western juniper site, where low carbon availability reduced assimilatory demands for nitrogen.

The high rates of microbial assimilation at these sites raise an important question regarding carbon availability: is sufficient carbon fixed in these ecosystems to sustain the high assimilation rates throughout the active season? Whether or not the rates shown in Table 2 can be sustained depends on: the length of the active period; the amount of above-ground and below-ground net primary production (ANPP and BNPP); the amount of NH_4^+ assimilated; the C:N ratio of the microbial biomass; and the carbon-use efficiency of the microbial biomass. We used a model describing secondary production processes¹⁶ to calculate

TABLE 1 Selected characteristics of study sites

Ecosystem	Symbol	Elevation (m)	MAT (°C)	MAP (mm)	ANPP* (Mg ha ⁻¹ yr ⁻¹)	Soil texture†	Soil pH‡	Extractable inorganic N§		¹⁵ NO ₃ ⁻ added (mg N kg ⁻¹)
								NH ₄ ⁺	NO ₃ ⁻	
Tesuque watershed, New Mexico										
Pinyon/juniper (<i>Pinus edulis</i> / <i>Juniperus osteosperma</i>)	Y	2,410	8.5	480	ND¶	SL	6.1	3.29	1.04	1.18
Ponderosa pine (<i>Pinus ponderosa</i>)	P	2,740	6.5	550	ND	SL	5.0	1.39	0.19	1.06
Mixed conifer (<i>Pseudotsuga menziesii</i> / <i>Abies concolor</i>)	C	2,720	5.0	550	ND	SL	5.1	2.65	0.23	1.47
Aspen (<i>Populus tremuloides</i>)	A	3,110	2.5	625	5.7	L	5.1	4.41	0.24	1.24
Spruce/fir (<i>Picea engelmannii</i> / <i>Abies lasiocarpa</i>)	S	3,415	0.0	700	ND	L	4.6	1.32	0.25	0.77
Oregon transect										
Western juniper (<i>Juniperus occidentalis</i>)	J	930	9.0	220	1.2	SL	5.9	1.32	0.50	1.02
Ponderosa pine (<i>Pinus ponderosa</i>)	O	1,030	7.5	540	2.2	SL	5.5	0.98	0.05	1.27
Mountain hemlock (<i>Tsuga mertensiana</i>)	M	1,460	6.0	1,800	5.1	SL	4.8	0.97	0.18	2.06
Douglas-fir (<i>Pseudotsuga menziesii</i>)	D	170	11.0	1,000	11.6	SiL	5.0	2.59	0.54	1.30
Red alder/Douglas-fir (<i>Alnus rubra</i> / <i>Pseudotsuga menziesii</i>)	R	200	10.0	2,400	10.7	L	3.5	4.11	1.24	3.23
Western hemlock/sitka spruce (<i>Tsuga heterophylla</i> / <i>Picea sitchensis</i>)	H	240	10.0	2,400	13.0	SiL	3.8	4.53	0.08	2.02

Soil characteristics are for the 0–15-cm mineral soil layer. ND, not determined; MAT, mean annual temperature; MAP, mean annual precipitation.

* Above-ground net primary production (in $\text{Mg biomass ha}^{-1} \text{ yr}^{-1}$) from refs 3, 11, 12, 29 and 30.

† L, loam; SL, sandy loam; SiL, silt loam.

‡ Soil pH was measured in 0.01 M CaCl_2 (2:1 solution:soil wt ratio).

§ Mean 2 M KCl-extractable N from spring and summer sampling dates; $n = 10$ for all sites, except for the spruce/fir and red alder/Douglas-fir sites, where $n = 5$.

the microbial carbon-use efficiencies that would be necessary to allow the nitrogen-assimilation rates to be sustained throughout the active periods at the sites along the Oregon transect. The length of the active period was estimated on the basis of the amount of time that the surface soil was moist but not frozen (periods estimated from data in refs 11 and 12 ranged from 96 to 335 days). Because no BNPP data are available for these sites, we assumed that BNPP was equal to ANPP (ratios of BNPP/ANPP estimated for forests typically range from 0.3 to 4, with the highest values obtained when the NPP translocated to mycorrhizal fungi is included in estimates^{17,18}). In addition, we summed mean NH_4^+ assimilation rates, measured in separate soil cores using $^{15}\text{NH}_4^+$ isotope dilution (data not shown), with mean NO_3^- assimilation rates for spring and summer to obtain rates of microbial $\text{NH}_4^+ + \text{NO}_3^-$ assimilation. We also assumed that two thirds of all microbial nitrogen assimilation occurs in the 0–15-cm soil layer. Finally, we used microbial C:N ratios that we had measured

following isotope dilution experiments at each site. Modelling results showed that assimilation rates could be sustained throughout the active periods if microbial carbon-use efficiencies at each site were as follows: juniper, 0.50; ponderosa pine, 0.33; mountain hemlock, 0.44; Douglas-fir 0.25; red alder/Douglas-fir, 0.57; and western hemlock/sitka spruce, 0.53. These carbon-use efficiencies are all within the range of values previously reported for soil microorganisms utilizing NO_3^- and NH_4^+ as N sources¹⁹.

The high rates of microbial assimilation of NO_3^- shown in these forests may have resulted from the activity of both saprophytic microorganisms and mycorrhizal fungi. In other coniferous forests of the northwestern coast of the United States, however, mycorrhizal fungi have been estimated to process <25% of the NPP^{17,20}, indicating that the major portion of the NPP is decomposed by saprophytic organisms. Therefore, the NO_3^- assimilation rates shown in Table 2 are most likely to be due to the activity of saprophytic microorganisms. It is also likely that the small amount of nitrogen assimilated by mycorrhizal fungi is used to produce fungal biomass rather than plant biomass. Plant demand for nitrogen, estimated on the basis of litterfall data^{3,11}, is more than an order of magnitude lower than the rates of NO_3^- assimilation that we measured.

The results of our study contradict two assumptions commonly made regarding nitrogen cycling in forest ecosystems: (1) that nitrification is insignificant in undisturbed coniferous forests; and (2) that microbial assimilation of NO_3^- is unimportant in controlling nitrogen retention in forests. Our results demonstrate that even in mature ecosystems with relatively 'tight' nutrient cycles²¹, large amounts of NO_3^- can be produced internally without subsequent nitrogen loss. In these systems, microbial assimilation of NO_3^- constitutes an important mechanism for nitrogen retention.

We propose a modification of the existing paradigm for nitrogen retention in forest ecosystems that accounts for high rates of microbial assimilation of NO_3^- . In ecosystems where soil carbon and nitrogen are accumulating, the microbial biomass will be a net sink for inorganic nitrogen, including NO_3^- . In mature ecosystems, however, where annual increases in microbial biomass are relatively small, microbial biomass nitrogen accretion is unlikely to cause rates of NO_3^- assimilation as high as those shown here. Instead, we propose that rapid turnover of the microbial biomass promotes nitrogen retention. Microbial assimilation of NO_3^-

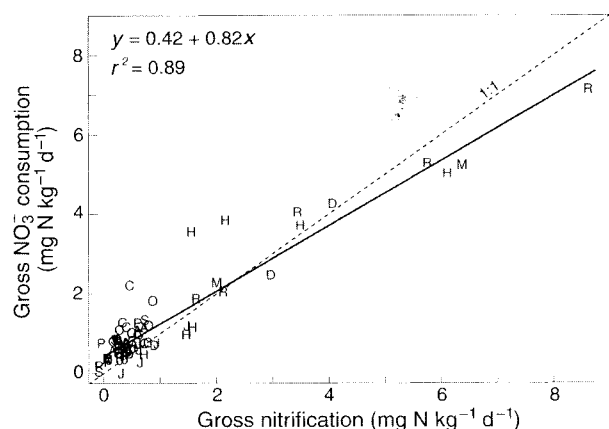


FIG. 1 Relationship between NO_3^- production and consumption in intact soil cores. Points represent gross rates in individual cores measured by ^{15}N -isotope dilution during late spring and late summer in the 0–15-cm mineral soil layer of forests in New Mexico and Oregon ($n = 70$). See Table 1 for key to symbols.

TABLE 2 Gross rates of nitrification and NO_3^- consumption, net rates of nitrification, and mean residence times for NO_3^- pools in mature forest ecosystems

Ecosystem	Gross nitrification ($\text{mg N m}^{-2} \text{ d}^{-1}$)		Gross NO_3^- consumption ($\text{mg N m}^{-2} \text{ d}^{-1}$)		Net NO_3^- accumulation ($\text{mg N m}^{-2} \text{ d}^{-1}$)		NO_3^- mean residence time (d)	
	Spring	Summer	Spring	Summer	Spring	Summer	Spring	Summer
Tesuque watershed, New Mexico								
Pinyon/juniper	39*	125 (33)*	28*	107 (20)*	-22 (48)*	66 (36)*	1.46*	0.27 (0.17)*
Ponderosa pine	29 (13)	25 (11)	60 (20)	80 (13)	-30 (12)	-54 (20)	1.08 (0.67)	0.88 (0.52)
Mixed conifer	62 (5)	44 (4)	86 (9)	164 (26)	-24 (11)	-120 (23)	0.42 (0.11)	0.45 (0.10)
Aspen	46 (7)	35 (10)	75 (10)	76 (11)	-29 (5)	-41 (2)	0.78 (0.48)	1.08 (0.40)
Spruce/fir	ND†	27 (13)	ND	66 (21)	ND	-39 (10)	ND	0.80 (0.10)
Oregon transect								
Western juniper	82 (18)	127 (37)	97 (14)	85 (36)	-15 (10)	42 (13)	1.47 (0.66)	1.31 (0.64)
Ponderosa pine	41 (22)	79 (21)	84 (44)	112 (37)	-58 (26)	-66 (21)	0.00 (0.00)	0.28 (0.25)
Mountain hemlock	ND	198 (119)	ND	189 (84)	ND	-65 (50)	ND	0.11 (0.04)
Douglas-fir	304 (132)	82 (25)*	304 (128)	42 (12)*	1 (29)	102 (53)*	1.16 (0.79)	0.34 (0.13)*
Red alder/Douglas-fir	270 (90)	ND	251 (70)	ND	19 (21)	ND	0.35 (0.11)	ND
Western hemlock/ sitka spruce	112 (12)	145 (54)	224 (5)	111 (53)	-109 (7)	34 (13)	0.07 (0.03)	0.00 (0.00)

Values are means and (standard errors); $n = 5$ for all net rates; $n = 5$ for gross rates and mean residence times at all sites, except: $n = 1$ where no standard error is listed; $n = 2$ at the western hemlock/sitka spruce site in spring and the Oregon ponderosa pine site in summer; $n = 3$ at the Oregon ponderosa pine site in spring and at the pinyon/juniper, Douglas-fir, and mountain hemlock sites in summer; and $n = 4$ at the western hemlock/sitka spruce site in summer. Gross nitrification rates were significantly greater than zero at all sites except the mountain hemlock site ($P = 0.12$ for the mountain hemlock site, $P \leq 0.10$ for the Oregon ponderosa pine site, and $P \leq 0.05$ for all other sites with $n > 1$).

* Rates could not be determined in intact cores because the soil was either too rocky or too hard; rates from homogenized samples are provided instead.

† Rates not determined.

followed by release of biomass nitrogen as organic nitrogen and NH_4^+ results in conversion of the highly mobile NO_3^- ion into less mobile nitrogen forms. Therefore, NO_3^- assimilation and rapid microbial turnover will continually deplete NO_3^- pools, and minimal NO_3^- leaching will occur in spite of significant rates of nitrification.

Although increased rates of nitrogen loss from ecosystems following disturbance are usually attributed to increased rates of nitrification, greater availability of NH_4^+ and reduced inputs of plant carbon following disturbance are likely to decrease microbial assimilation of NO_3^- . Therefore, the increased nitrogen loss observed in a large number of forest ecosystems following vegetation removal¹⁰ and increased nitrogen deposition²² may be largely due to a reduction in NO_3^- assimilation by soil microorganisms. □

Methods

On each sampling date, $^{15}\text{NO}_3^-$ isotope dilution measurements¹³ were made in intact soil cores at five plots distributed over 0.3 ha at each site. Each intact mineral soil core (4.8 cm diameter by 15 cm deep) received eight 2-ml injections of a solution containing 1.7 mM K^{15}NO_3 (99 atom% ^{15}N), resulting in an increase of 30 to 150 g $\text{H}_2\text{O kg}^{-1}$ oven dry soil and 0.77 to 3.23 mg $\text{NO}_3^- \text{N kg}^{-1}$ (Table 1), depending on soil bulk density. After injection of two core samples from each plot, one sample was immediately homogenized and a subsample was extracted in 2 M KCl to determine ^{15}N extraction efficiency¹³. Intact soil cores and homogenized samples were sealed in separate one-litre canning jars. Acetylene was added to half of the jars containing homogenized samples (creating 10 kPa C_2H_2). The jars were buried at the original location and soil depth for 24 h. After incubation, head-space samples were collected for N_2O analysis²³, and soil subsamples were analysed for NO_3^- (refs 24, 25), microbial biomass ^{15}N (refs 26, 27), and total organic ^{15}N (ref. 28). All analyses of microbial biomass ^{15}N and total organic ^{15}N included procedures for eliminating NO_3^- to prevent contamination of organic N by residual $^{15}\text{NO}_3^-$. Gross nitrification and NO_3^- consumption rates were calculated from isotope dilution equations¹³. Mean residence times were calculated assuming that ambient NO_3^- pools were at steady state and that fluxes were equal to gross nitrification rates.

Received 2 May; accepted 8 November 1996.

- Johnson, D. W. *J. Environ. Qual.* **21**, 1–12 (1992).
- Gosz, J. R. in *Terrestrial Nitrogen Cycles* (eds Clark, F. E. & Rossow, T.) 405–426 (Ecol. Bull., Stockholm, 1981).
- Vitousek, P. M., Gosz, J. M., Grier, C. C., Melillo, J. M. & Reiners, W. A. *Ecol. Monogr.* **52**, 155–177 (1982).
- Gosz, J. R. & White, C. S. *Biogeochemistry* **2**, 281–297 (1986).
- Davidson, E. A., Hart, S. C. & Firestone, M. K. *Ecology* **73**, 1148–1156 (1992).
- Jansson, S. L., Hallam, M. J. & Bartholomew, W. V. *Plant Soil* **6**, 382–390 (1955).
- Tiedje, J. M., Sorenson, J. & Chang, Y. Y. L. in *Terrestrial Nitrogen Cycles* (eds Clark, F. E. & Rossow, T.) 331–342 (Ecol. Bull., Stockholm, 1981).
- Myrold, D. D. & Tiedje, J. M. *Soil Biol. Biochem.* **18**, 559–568 (1986).
- Gosz, J. R. *Water Res.* **12**, 725–734 (1978).
- Vitousek, P. M. *et al. Science* **204**, 469–474 (1979).
- Gholz, H. L. *Ecology* **63**, 469–481 (1982).
- Rumyon, J., Waring, R. H., Goward, S. N. & Welles, J. M. *Ecol. Appl.* **4**, 211–225 (1994).
- Hart, S. C., Stark, J. M., Davidson, E. A. & Firestone, M. K. in *Methods of Soil Analysis* (eds Weaver, R. et al.) 985–1018 (Soil Sci. Soc. of America, Madison, Wisconsin, 1993).
- NADP/NTN Annual Data Summary 1983–1993 (Nat. Atmospheric Deposition Progm, Colorado State Univ., Fort Collins, Colorado, 1993).
- Binkley, D. & Hart, S. C. *Adv. Soil Sci.* **10**, 57–112 (1989).
- Strayer, D. *Limnol. Oceanogr.* **33**, 1217–1220 (1988).
- Vogt, K. A., Grier, C. C., Meier, C. E. & Edmonds, R. L. *Ecology* **63**, 370–380 (1982).
- Fogel, R. & Hunt, G. *Can. J. For. Res.* **13**, 219–232 (1983).
- Hart, S. C., Nason, G. E., Myrold, D. D. & Perry, D. A. *Ecology* **75**, 880–891 (1994).
- Fogel, R. & Hunt, G. *Can. J. For. Res.* **9**, 245–256 (1979).
- Odum, E. P. *Science* **164**, 262–270 (1969).
- Aber, J. D. *Trends Ecol. Evol.* **7**, 220–223 (1992).
- Mosier, A. R. & Klemedtsson, L. in *Methods of Soil Analysis*, Vol. 2 (eds Weaver, R. et al.) 1047–1065 (Soil Sci. Soc. of America, Madison, Wisconsin, 1994).
- Keeney, D. R. & Nelson, D. W. in *Methods of Soil Analysis*, Vol. 2, 2nd edn (eds Page, A. L., Miller, R. H. & Keeney, D. R.) 643–698 (Am. Soc. of Agron., Madison, Wisconsin, 1982).
- Stark, J. M. & Hart, S. C. *Soil Sci. Soc. Am. J.* **60**, 1846–1855 (1996).
- Brookes, P. C., Landman, A., Pruden, G. & Jenkinson, D. S. *Soil Biol. Biochem.* **17**, 837–842 (1985).
- Davidson, E. A., Eckert, R. W., Hart, S. C. & Firestone, M. K. *Soil Biol. Biochem.* **21**, 773–778 (1989).
- Barrie, A., Davies, J. E., Park, A. J. & Workman, C. T. *Spectroscopy* **4**, 44–52 (1989).
- Gosz, J. R. *Ecology* **61**, 507–514 (1980).
- Binkley, D., Sollins, P., Bell, R., Sachs, D. & Myrold, D. *Ecology* **73**, 2022–2033 (1992).

ACKNOWLEDGEMENTS. We thank K. Haubensack for help with data collection and analysis; G. Mileski Y. Myint, E. Smith, J. Hart, R. Wright, P. Roosma, D. Smart, M. Swarts, V. Diego and J. Wyllie for assistance with field and laboratory work; C. Glassman for additional laboratory facilities; C. S. White and R. H. Waring for help in identifying study sites; and J. D. Aber and A. Stark for comments on the manuscript. This research was supported by grants from the US NSF and USDA–CSRS.

CORRESPONDENCE and requests for materials should be addressed to J.M.S. (e-mail: jstark@cc.usu.edu).

The *Drosophila* protein Wunen repels migrating germ cells

Nian Zhang^{*†}, Jiaping Zhang^{†‡}, Karen J. Purcell[§], Yan Cheng^{†||} & Ken Howard^{†§¶}

^{*} The Jackson Laboratory, 600 Main Street, Bar Harbor, Maine 04609, USA

[†] Roche Institute of Molecular Biology, Roche Research Center, 340 Kingsland Street, Nutley, New Jersey 07110, USA

[‡] Department of Virus and Cell Biology, Merck Research Laboratories, West Point, Pennsylvania 19486, USA

[§] Boyer Center for Molecular Medicine, Yale University School of Medicine, Departments of Cell Biology and Biology, 295 Congress Avenue, P.O. Box 9812, New Haven, Connecticut 06536-0812, USA

^{||} Department of Biological Sciences, University of Notre Dame, Notre Dame, Indiana 46556, USA

[¶] Laboratory for Molecular Cell Biology, University College London, Gower Street, London WC1E 6BT, UK

IN *Drosophila*, germ cells migrate in embryonic development from the lumen of the developing gut towards the overlying mesoderm, where they enter the gonads^{1,2}. The gene *wunen* is responsible for guiding the germ cells early in this process³. Here we report that the protein Wunen has two properties that allow it to use repulsion to guide the germ cells. Wunen can transform a permissive cellular environment into a repulsive one, and is expressed in the gut in a pattern that guides germ cells towards the mesoderm. Wunen shows strong similarity to the enzyme type 2 phosphatidic acid phosphatase (PAP2)⁴, suggesting that it is involved in lipid metabolism.

We have previously mapped the gene *wunen* (*wun*) to 45D on the second chromosome³, and subsequently identified a *wun* allele, *wun*^{K10201}, induced by a marked P element from a large-scale screen⁵. Excision of this P element reverts *wun*^{K10201}, indicating it is the *wun* mutagen in this allele. We cloned the DNA flanking this P element, localized these sequences in the contig we had previously isolated³, and identified transcripts in this region (Fig. 1). These transcripts, arbitrarily designated A, B and C, encode a gene with homology to the ABC transporter family⁶ (transcript A), one with sequence similarity to the zinc-finger transcription factor GFI-1 (ref. 7) (transcript B), and one homologous to PAP2 (transcript C); the sequence of transcript C is shown in Fig. 2.

Transcript A, which shows uniform expression in embryos, is downregulated in the *wun*⁺ mutant *l(2)03659*⁰³⁶⁵⁹ (data not shown). This means that A cannot be the source of Wunen (Wun) activity, and left us to distinguish between transcripts B and C. The P element in *wun*^{K10201} truncates the product of transcript B and downregulates the expression of transcript C, effectively abolishing both functions. The transformation construct NX restores the coding sequence of transcript B and rescues the lethality of *wun*^{K10201} without restoring germ-cell migration, suggesting that transcript B cannot be the source of *wun* function. The original *wun* allele, *wun*^{CE}, also downregulates the expression

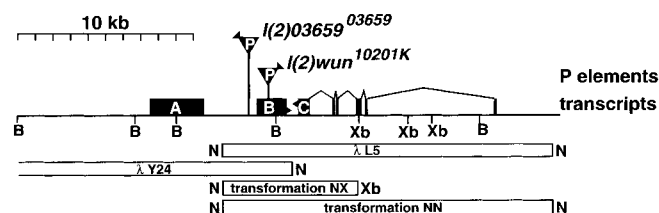


FIG. 1 DNA map. B, BamHI; Xb, XbaI; N, NotI. Scale is in kilobases. The two P-element inserts are shown as triangles above the map and are in different orientations. The orientation of transcription for B and C is shown by arrows on the transcript; orientation was not determined for A.

of transcript C (data not shown), further suggesting that transcript C encodes *wun* function. Support for this idea also comes from the fact that the larger transformation construct NN, which supplies both transcripts B and C, rescues both the lethality and the migration defect of *wun*^{K10201} and *wun*^{CE}. Although these data do not in themselves demonstrate that transcript C encodes the *wun* function, we obtained data on the expression pattern and activity of this product that argue very strongly that this is the case.

This product is first expressed at high levels in the hindgut primordium, the distal part of the gut where germ cells are not normally found but can be found in the *wun* mutant (Fig. 3e). As the germ cells leave the primordial gut and move away from its lower side towards the overlying mesoderm, transcript C is expressed specifically on the lower side of this tissue (Fig. 3f). In the *wun* mutant this orientation is abolished (compare the wild type shown in Fig. 3a with the *wun* mutant in Fig. 3b).

To determine whether the product encoded by transcript C could inhibit germ-cell migration, we overexpressed it in the mesoderm, which does not normally express *wun* and supports germ-cell migration to the gonad. If we express *wun* here, very few germ cells enter the mesoderm, most remaining in the yolk (Fig. 3i, j). This phenotype is remarkably different from that caused by lack of *wun* function, which abolishes guidance, allowing the germ

cells to move largely at random. Mesodermal overexpression of *wun* does not prevent development of the gonadal mesoderm, which forms relatively normally (Fig. 3k, l). We conclude that *wun* encodes a specific component of a repulsive-guidance system for germ cells, which first acts as the germ cells interact with the gut, and that this function is provided by transcript C.

The product encoded by transcript C is a protein of 300 amino acids, and is a member of a newly identified protein family (Fig. 2). The completely sequenced members of this group, identified in genomic sequencing projects and in an experimental context as a gene encoding PAP2 (ref. 4), are shown aligned in Fig. 2a. The predicted mature *Drosophila* protein, Wun, is shown wound through the membrane in Fig. 2d. Note the highly conserved residues in the intercellular space which might contribute to the active site of this class of molecules. The extracellular domains of these proteins are much less highly conserved. We suggest that they mediate regulatory interactions, rendering such molecules receptors which feed directly into the production of diacyl glycerol. The literature in this field suggests several possible pathways to investigate.

In the extended germ band, *wun* is expressed in the ectoderm in a medial band throughout the trunk (Fig. 3i). It might be required to drive the germ cells to move bilaterally as they enter the

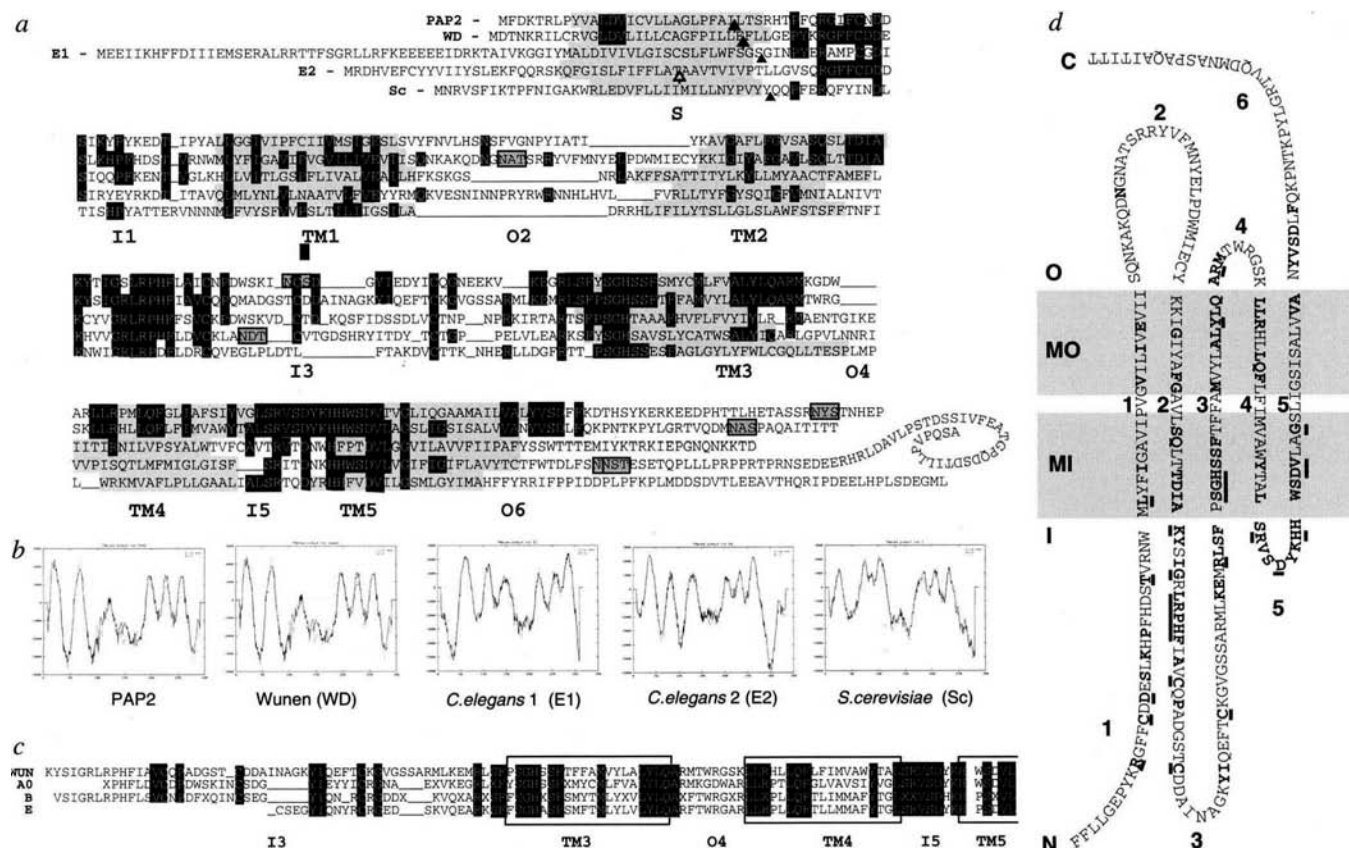


Fig. 2 Sequences. a, Alignment of the sequences discussed here. WD refers to *Drosophila* Wun. Sequence E1 is present in the *Caenorhabditis elegans* cosmid with accession number F13E6; E2 with U28738 and the *Saccharomyces cerevisiae* (Sc) sequence has accession number U51031. Conserved residues are white with a black background. Transmembrane domains are shaded light grey with no border and named S (putative signal sequence) and TM1 to TM5. The locations of putative signal sequence cleavage sites are marked with arrows, and consensus glycosylation sites indicated with shaded boxes with black borders. The orientation in the membrane was predicted independently for all five examples and is

consistent with the location of glycosylation sites in PAP2, which is glycosylated by experimental criteria⁸. b, Graphical output of the TM predict routine for these five sequences. c, Homologues identified in expressed sequence tags with accession numbers and sources of RNA in parentheses: A0, n310479 (human melanocyte) and r97295 (human fetal liver/spleen); B, r63796, r71019, h04659 (all human placenta) and t92854 (human lung); E, r75377 (mouse brain). d, The *Drosophila* Wun sequence wound through the membrane (MI, inner leaflet; MO, outer leaflet). Residues conserved between PAP2 and Wun are marked in bold type, and those which are also shared with *C. elegans* are underlined.

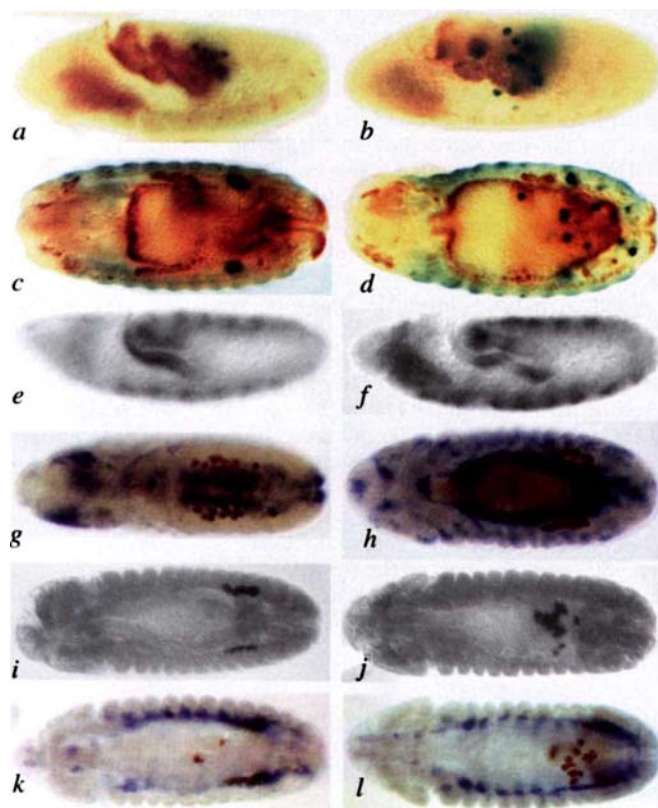


FIG. 3 Phenotypes. Germ cells (blue) and gut (brown) in a wild-type (a) and a *wun* mutant (b) embryo at stage 11. These cells remain dispersed in germ-band-shortened *wun* mutants (d), but a gonad forms in the wild type (c). e–h, Expression of *wun* RNA. e, The *wun* transcript is detected in the hindgut, and a segmental repeat is found in the ectoderm. f, It is later expressed strongly on the dorsal (embryonic axes) side of the posterior midgut. g, Expression of *wun* (blue) in a medial ectodermal stripe as the germ cells (brown) move bilaterally in the mesoderm. h, Expression of *wun* in the gut, nervous system and ectoderm of germ-band-shortened embryo. i–l, Control and experimental embryos from the overexpression experiment. i, A wild-type control. j, An experimental embryo where germ cells do not enter the mesoderm and remain in the interior of the embryo. k, l, Two embryos like those shown in i and j, showing the somatic gonad (blue) and germ cells (brown).

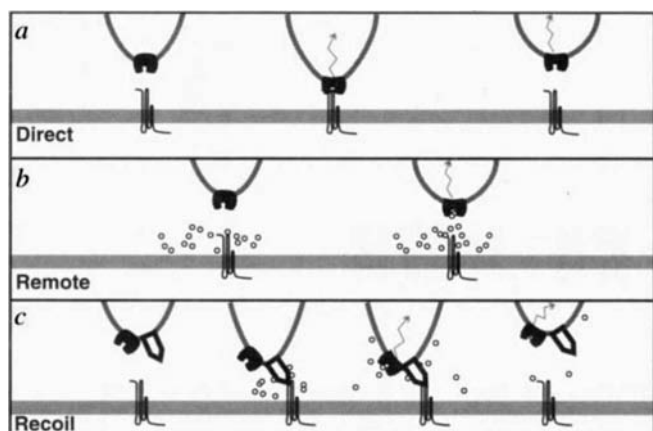


FIG. 4 Models. a, In the direct model, contact between the cells is required for signalling which occurs with the extracellular domain of Wun acting as a membrane-bound ligand. b, In the remote model, Wun constitutively releases a diffusible repellent species from cells where it is expressed. c, In the recoil model, contact between a germ cell and a Wun-expressing cell activates the enzymatic activity of Wun and releases the signal.

mesoderm, and may explain the later dispersion seen in *wun* mutants. However, we think this is unlikely and, based on the consequences on germ-cell migration of the absence of *nanos*⁸, we suggest that the *wun* signal be required for the germ cells to differentiate properly and detect later guidance signals. In addition, the role of the other, later expression of *wun* (Fig. 3g, h) is not yet clear. We have studied another migratory population, neurons, and have not seen defects in *wun* mutants.

Guidance mechanisms that use repulsion are seen in the development of the nervous system where the molecularly characterized protein factors, Unc-6/netrin^{9–11}, collapsin^{12,13} and RAGS¹⁴ repel growth cones from certain classes of neurons. These molecules all act as signal-transducing ligands that impinge on the responding cells through appropriate receptors. We consider it unlikely that the extracellular domain of Wun acts as a membrane-bound ligand, like the *Drosophila* protein Bride of sevenless¹⁵ (Fig. 4a). Instead, we expect Wun to have enzymatic activity; it might be constitutively active, causing all the tissues where it is expressed to release a repellent (Fig. 4b). We favour a slightly more complex model in which contacts^{16,17} between germ cells and Wun-expressing cells activate Wun and release the repulsive signal. This 'recoil' model is shown in Fig. 4c. In the extreme, this model allows a single repulsive species to be involved in many different migrations by tight regulation of its production, which is stimulated by specific contact events.

There are many unresolved questions about Wun, including: what enzymatic activity it shows; whether this activity is regulated; how Wun acts to produce a repulsive signal; and whether any other cells respond to this signal. There are also broader questions about the roles of other members of this family. These questions can be addressed by a combination of genetics and cell biology in *Drosophila* and other systems.

Note added in proof: I. Wada and H. Kanoh have pointed out that the N terminus of PAP2 is not cleaved, and that WUNEN probably spans the membrane six, rather than five, times. □

Methods

Immunocytochemistry, *in situ* hybridization, and molecular cloning procedures were by standard methods. The embryos were stained to reveal one or two of the following: germ cells using anti-vasa¹⁸ antibody; somatic gonad using anti 412 (ref. 19) *in situ* hybridization; *wun* using anti-transcript C *in situ* hybridization; or the gut tissue using anti-β-galactosidase antibody to the enhancer trap line 4850. Transcripts in the region were identified by *in situ* hybridization to embryos and by probing northern blots with DNA fragments covering all of the insert of λL5. cDNA clones were then isolated from libraries prepared in λlocus by B. Morris (Novagen). Genes similar to *wun* were identified using the BLAST algorithms²⁰ run at the NCBI (<http://www.ncbi.nlm.nih.gov/>) and the Sanger Center (<http://www.sanger.ac.uk/>). Alignment was aided using the e-mail server at cbrg@inf.ethz.ch. Transmembrane domains were predicted using TM predict²¹ (<http://ulrec3.unil.ch/software/>). The locations of putative signal sequence cleavage were indicated by PSORT (<http://psort.nibb.ac.jp/>). The overexpression construct contained the *wun* cDNA c18 cloned into the EcoRI and Asp 718 sites of pUAST²². These flies were crossed to the twist-GAL-4 line²³, causing expression of *wun* in the mesoderm using GAL4 activation of transcription from the UAS sites in pUAST. Embryos were collected and stained to reveal germ cells and the somatic gonad. Two different transformant lines produced indistinguishable results.

Received 5 September; accepted 30 October 1996.

1. Sonnenblick, B. P. *Proc. Natl. Acad. Sci. USA* **27**, 484–489 (1941).
2. Rabinowitz, M. J. *Morphol.* **69**, 1–49 (1941).
3. Zhang, N., Zhang, J.-P., Cheng, Y. & Howard, K. R. *Genetics* **143**, 1231–1241 (1996).
4. Kai, M., Wada, I., Imai, S., Sakane, F. & Kanoh, H. *J. Biol. Chem.* **271**, 18931–18938 (1996).
5. Torok, T., Tick, G., Alvarado, M. & Kiss, I. *Genetics* **135**, 71–80 (1993).
6. Faith, M. J. & Kotler, R. *Microbiol. Rev.* **57**, 995–1017 (1993).
7. Gilks, C. B., Bear, S. E., Grimes, H. L. & Tschlis, P. N. *Mol. Cell. Biol.* **13**, 1759–1768 (1993).
8. Kobayashi, S., Yamada, M., Asaoka, M. & Kitamura, T. *Nature* **380**, 708–711 (1996).
9. Hedgecock, E. M., Culotti, J. G., Hall, D. H. & Stern, B. *Development* **100**, 365–382 (1987).
10. Ishii, N., Wadsworth, W. G., Stern, B. D., Culotti, J. G. & Hedgecock, E. M. *Neuron* **9**, 873–881 (1992).
11. Colamarino, S. A. & Tessier, L. M. *Cell* **81**, 621–629 (1995).
12. Luo, Y., Raible, D. & Raper, J. A. *Cell* **75**, 217–227 (1993).
13. Fan, J., Mansfield, S. G., Redmond, T., Gordon, W. P. R. & Raper, J. A. *J. Cell Biol.* **121**, 867–878 (1993).
14. Drescher, U. et al. *Cell* **82**, 359–370 (1995).
15. Hart, A. C., Kramer, H. & Zipursky, S. L. *Nature* **361**, 732–736 (1993).

16. Callaini, G., Giovanna Ripabelli, M. & Dalli, R. *Dev. Biol.* **170**, 365–375 (1995).
17. Jaglarz, M. K. & Howard, K. R. *Development* **121**, 3495–3504 (1995).
18. Hay, B., Jan, L. Y. & Jan, Y. N. *Cell* **55**, 577–587 (1988).
19. Brookman, J. J., Toosy, A. T., Shashidhara, L. S. & White, R. A. H. *Development* **116**, 1185–1192 (1992).
20. Altschul, S. F., Gish, W., Miller, W., Myer, E. W. & Lipman, D. J. *J. Mol. Biol.* **215**, 403–410 (1990).
21. Hofmann, K. & Stoffel, W. *Hoppe-Seyler's Z. Biol. Chem.* **347**, 166 (1993).
22. Brand, A. H. & Perrimon, N. *Development* **118**, 401–415 (1993).
23. Grieg, S. & Akam, M. *Curr. Biol.* **5**, 1057–1062 (1995).

ACKNOWLEDGEMENTS. We thank Hoffmann La Roche for supporting this work at the Roche Institute of Molecular Biology in Nutley, New Jersey, and S. Artavanis-Tsokanos for hosting K.H. at Yale. We thank M. Akam, A. Brand, H. Broihier, S. Cockcroft, M. Jaglarz, A. Hayden, T. Jessell, I. Kiss, S. Kobayashi, T. Lavery, R. Lehmann, B. Morris, C. Rushlow, H. Skaer, C. Stewart, M. van Dorean, P. Wassermann and A. Williamson for discussion and materials.

CORRESPONDENCE and requests for materials should be addressed to K.H. (e-mail: yumi@panix.com). The sequences have been deposited at Genbank, accession numbers U73821, U73822 and U7823.

Regulation of neuronal diversity in the *Xenopus* retina by Delta signalling

Richard I. Dorsky*, Wesley S. Chang*, David H. Rapaport† & William A. Harris*

Department of *Biology, †Division of Anatomy, Department of Surgery, University of California, San Diego, 9500 Gilman Drive, La Jolla, California 92093, USA

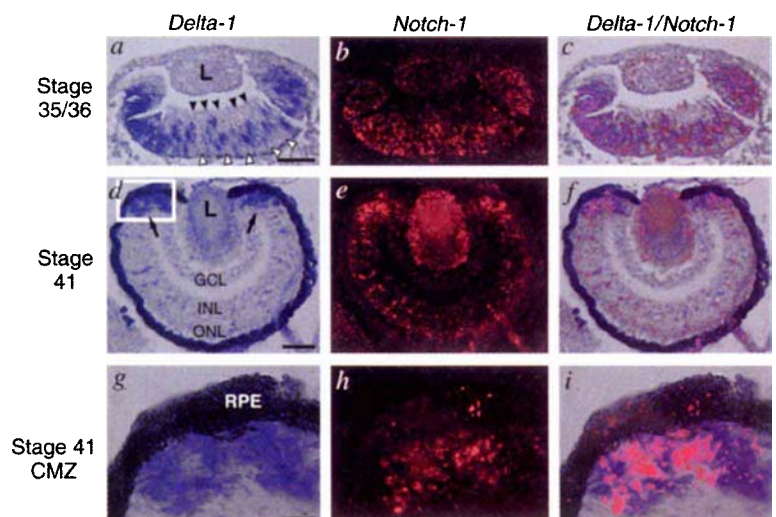
To generate the variety of mature neurons and glia found in the developing retina, the competence of pluripotent progenitor cells to respond to extracellular signals must be controlled. Delta, a ligand of the Notch receptor, is a candidate for regulating progenitor competence on the grounds that activation of the pathway involving Notch and Delta can inhibit cellular differentiation^{1–6}. Here we test this possibility in the developing *Xenopus* retina by misexpression of Delta messenger RNA. We find that Delta-misexpressing cells with wild-type neighbours adopt earlier fates, primarily becoming ganglion cells and cone photoreceptors. Progenitors transfected with Delta later in development also produce rod photoreceptors, but not the latest-generated cell types, demonstrating the importance of timing in Delta function. We conclude that Delta signalling in the vertebrate retina is a basic regulatory mechanism that can be used to generate neuronal diversity.

FIG. 1 Retinal expression patterns of *X-Delta-1*, *X-Notch-1* and overlapping regions of expression. *X-Delta-1* is shown in blue, *X-Notch-1* in red, and areas of co-expression are pink. a–c, At stage 35/36, both genes are expressed in the immature cells of the central retina, but not where mature ganglion cells (black arrowheads) or photoreceptors (white arrowheads) are located. Expression is highest in the peripheral retina. d–f, By stage 41, expression is limited to the ciliary marginal zone (arrows) and late-born cells in the central retina. g–i, In the ciliary marginal zone, most cells express both genes. The white box in d indicates the area shown in g–i. L, lens; GCL, ganglion cell layer; INL, inner nuclear layer; ONL, outer nuclear layer; RPE, retinal pigment epithelium. Scale bars: a, d, 100 µm; g, 20 µm.

Cell types in the vertebrate retina are generated in a conserved temporal order, with retinal ganglion, horizontal and cone cells being born first, followed by amacrine, rod, bipolar and Müller glial cells^{7–9}. Clonal analysis, however, has shown that retinal fate is established by a lineage-independent mechanism^{7,10,11}. Furthermore, short-range signals are required to induce specific retinal cell types^{12–15}. Growth factors and hormones such as TGF- α /EGF, CNTF, retinoic acid and thyroxine can influence the number of cells that become rods and cones^{16–19}. A critical question underlying these results is why all cells do not respond to these signals, creating a uniform pattern of neuronal determination. The answer may lie in the competence of cells to process inductive cues, a mechanism that is controlled by the Delta–Notch lateral-inhibitory signalling pathway²⁰. In the *Xenopus* retina, the last cells to express *X-Notch-1* assume the last-generated fate, that of Müller glia, and cells expressing a dominant-activated form of *X-Notch-1* do not differentiate¹. This, and other results from *Drosophila*, suggests that to generate cellular diversity both in the retina and elsewhere, ligands for Notch must create temporally and spatially restricted patterns of receptor activation, which in turn inhibit cells from differentiating in response to local inductive cues^{2,3,21,31}. One of the key features of the lateral inhibitory model in *Drosophila* is that activation of Notch downregulates Delta expression in the same cell. This signal is fed back as a lack of Notch activation on Delta-expressing neighbours, accentuating the different levels of Notch activity between cells. This model predicts that Delta misexpression should lead to a decrease in Delta expression in neighbouring cells and accelerated differentiation compared with these neighbours²⁰.

The gene encoding Delta, a ligand that activates Notch-mediated inhibition in the late *Xenopus* gastrula^{6,22}, is strongly expressed in the developing retina, but as soon as the first cells become postmitotic, its expression becomes restricted. Thus, by stage 35/36 (ref. 23) *X-Delta-1* expression in the photoreceptor and ganglion cell layers is decreased (Fig. 1a). By stage 41, when the retina is completely laminated, *X-Delta-1* expression is limited to the proliferative ciliary marginal zone and the latest-born cells of the central retina (Fig. 1d). This pattern mimics *X-Notch-1* staining in the *Xenopus* retina¹, and double *in situ* hybridization, used to determine the extent of overlap of expression between the two genes, showed that both messenger RNAs were localized to the same populations of cells (Fig. 1). The coordinated expression of these two genes is consistent with possible ligand–receptor signalling in retinal development.

X-Delta-1 mRNA was injected into 16-cell *Xenopus* embryos, targeting the blastomere which produces over half of the ipsilateral retina²⁴. Clones of misexpressing cells at stage 31 were



identified by immunostaining, and the wild-type neighbouring cells expressing *X-Delta-1* mRNA were counted. Only $7 \pm 3\%$ (s.e.m.; 5 animals) of these clones were surrounded by immediate neighbours expressing *X-Delta-1* at a level comparable to other wild-type cells expressing the gene in the same section (Fig. 2a, b). In contrast, $90 \pm 4\%$ (s.e.m.; 5 animals) of clones misexpressing a control protein, β -galactosidase, were neighbours to *X-Delta-1*-expressing cells. These results support the idea that a high concentration of Delta protein inhibits expression of the same gene in surrounding cells, as predicted by a lateral inhibitory model²⁰.

At stage 33/34, Delta-misexpressing progenitors showed a bias towards early differentiation. In clones misexpressing green fluorescent protein (GFP), $75 \pm 6\%$ (s.e.m.; 8 animals) of cells were neuroepithelial, with processes spanning the retina and cell bodies located in the presumptive inner nuclear layer. However, in small clones (<10 cells), only $7 \pm 2\%$ (s.e.m.; 11 animals) of Delta-

misexpressing cells surrounded by wild-type neighbours exhibited neuroepithelial morphology (Fig. 2c, d). Inside large clones (>25 cells), Delta-misexpressing cells were neuroepithelial (Fig. 2e), whereas those bordered by wild-type cells were not. The early differentiation of Delta-misexpressing cells with wild-type neighbours suggests that there is a lack of Delta signalling back onto the misexpressing cells.

By stage 41, over 85% of progenitors misexpressing Delta in small clones became cones, ganglion and horizontal cells, the earliest-generated cell types of the retina^{8,9} (Figs 2f, 2g, 3a). In these clones, the number of progenitors becoming any of the later-generated cell types (bipolar, amacrine, rod or Müller cells) was significantly decreased (Fig. 3a). Delta-misexpressing cells in the centre of large clones, however, were neuroepithelial in morphology. Very few of these cells were immunopositive for the protein islet, a ganglion cell marker, or for calbindin, a cone marker (Fig. 2h and Table 1), and none was immunopositive for later cell-type markers, including rod opsin and R5 (Müller cells). In contrast, within one cell-diameter of the border of these clones, most cells stained for islet or calbindin (Fig. 2h and Table 1). Delta^{STU}, a truncated form of the ligand, behaves as an antimorph in the *Xenopus* neural plate, leading to early differentiation and increasing the number of primary neurons, by interfering with activation of the Notch pathway⁶. In all sizes of clones misexpressing Delta^{STU}, over 95% of cells were immunopositive for either islet

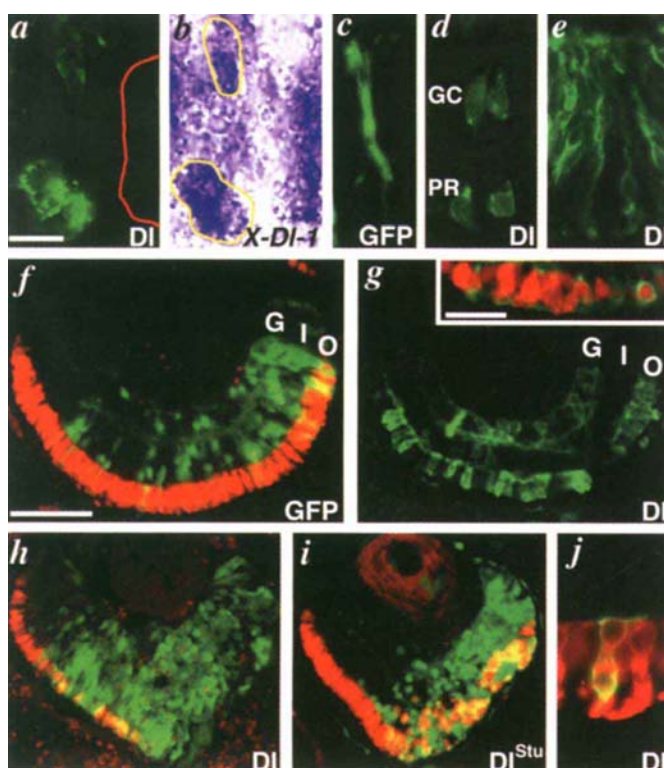


FIG. 2 Effects of Delta misexpression in the retina. a, After mRNA injection, Delta-misexpressing cells (green) are found in clones of cells at stage 31. b, *X-Delta-1* gene expression (blue) is detectable in the same clones (outlined in yellow), but neighbouring cells show only low levels of endogenous gene expression. Wild-type cells more than one cell-diameter away (red outline in a) show high levels of *X-Delta-1* expression. c, A small clone misexpressing GFP at stage 33/34 contains neuroepithelial cells. d, Cells in a clone misexpressing Delta have begun to differentiate into photoreceptors (PR) and ganglion cells (GC) by the same stage. e, A large clone misexpressing Delta contains neuroepithelial cells internally. f, At stage 41, GFP-misexpressing cells from radial clones were comprised of all cell types. Calbindin staining (cones) is in red. g, In contrast, Delta-misexpressing cells (green) are located almost exclusively in the GCL (G) and ONL (O), but not in the INL (I), and labelled photoreceptors are calbindin-positive cones (red, inset). h, In the centre of a large clone, cells misexpressing GFP and Delta (green) are neuroepithelial in morphology and immunonegative for calbindin (red); normal calbindin staining is seen at the border. i, In contrast, cells coexpressing GFP and Delta^{STU} are rounded like mature neurons and most express calbindin (red) or islet. j, A rod photoreceptor transfected with *X-Delta-1* cDNA at stage 18 (green) is labelled with a rod opsin antibody (red). In all panels, colocalization of fluorescent staining appears yellow. Scale bars: a-e, j, 10 µm; f-i, 50 µm; g (inset), 20 µm.

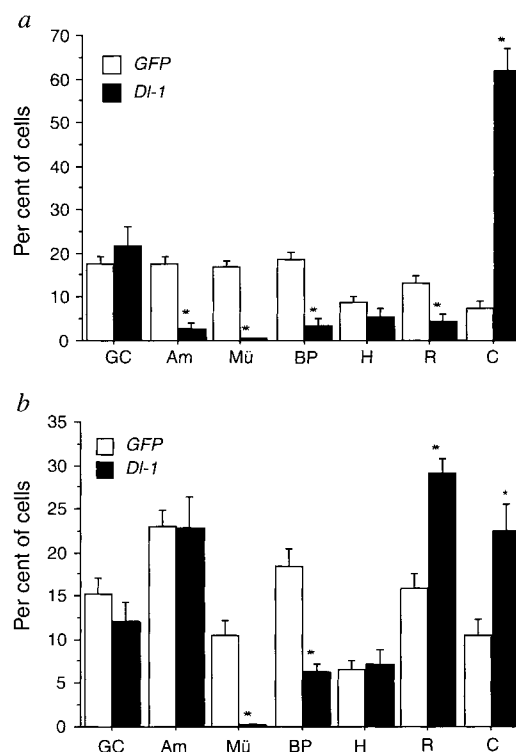


FIG. 3 Percentages of cell types produced by misexpression experiments, in clones of less than 10 misexpressing cells. a, Cells produced by mRNA injections into a single blastomere at the 16-cell stage. *X-Delta-1* injections produced significantly more cones and significantly fewer amacrine, Müller, bipolar and rod cells than GFP-injected controls. Ganglion and horizontal cell numbers were not significantly affected. $n = 1,201$ cells in 7 animals (GFP) and 663 cells in 12 animals (*X-Delta-1*). b, Cells produced by DNA transfection at stage 18. *X-Delta-1* transfection produced significantly more rods and cones and significantly fewer bipolar cells and Müller glia than GFP transfections. $n = 751$ cells in 9 animals (GFP) and 312 cells in 5 animals (*X-Delta-1*). Error represents s.e.m.; * $P < 0.01$, ** $P < 0.001$ by Student's *t*-test. GC, ganglion cell layer cells (includes displaced amacrine cells); Am, amacrine cells; Mü, Müller glia; BP, bipolar cells; H, horizontal cells; R, rods; C, cones.

or calbindin (Fig. 2i and Table 1). There were no cells positive for rod or Müller cell markers in Delta^{STU}-expressing clones. These results are consistent with the idea that the cells freed from lateral inhibition become competent to adopt the first available fates induced by their environment.

Notch activity in *Drosophila* affects different cell-fate decisions at different times in development^{5,25}. To investigate whether this occurs in the vertebrate retina, progenitors were transfected with *X-Delta-1* DNA at stage 18. The earliest detectable expression was 6–8 hours later at stage 25, by which time the first retinal cells are postmitotic⁷. Because of the largely overlapping genesis of different cell types in the rapidly developing *Xenopus* retina, transfection at this stage is likely to affect all cell decisions, rather than just the early ones. In these experiments, over 50% of cells misexpressing Delta were rod or cone photoreceptors (Figs 2j, 3b). The number of *X-Delta-1*-transfected precursors developing into either type of photoreceptor was significantly greater than in controls expressing GFP. In addition, significantly fewer *X-Delta-1*-transfected cells became bipolar neurons or Müller glia than did controls (Fig. 3b). These experiments show that in a later environment, Delta-misexpressing progenitors adopt rod as well as cone fates at the expense of the latest-born cells in the *Xenopus* retina¹.

Retinal progenitors in different environments may have different developmental potentials: for example, early retinal progenitors, when isolated, can form both ganglion cells^{3,26} and cones²⁷ *in vitro*, and rat retinal progenitors differentiate primarily as ganglion cells when isolated at embryonic day 14, but produce rods when isolated at postnatal day 1 (ref. 26). When rat embryonic cells are mixed with an excess of postnatal cells, the younger cells assume rod instead of ganglion cell fates¹⁴. These results indicate that mature cells affect fate decisions of later-differentiating cells *in vivo*. One explanation for the specific increase in photoreceptors found here may be that they are induced by other retinal neurons, such as ganglion cells, whose numbers can be regulated by feedback mechanisms^{28,29}. Our results support a model in which

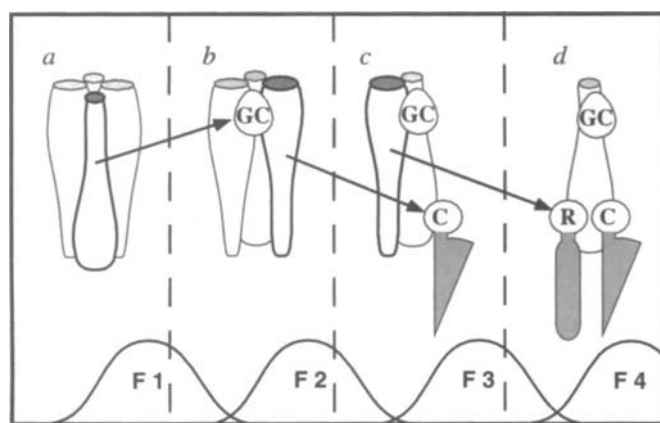


FIG. 4 Model of Delta action in vertebrate retinogenesis. *a*, Early in development, some neuroepithelial progenitors express high levels of Delta (dark shading), inhibit their neighbours from differentiating, and adopt early fates, becoming ganglion cells (GC). *b*, Once a photoreceptor-inducing signal is present in the retina, Delta-expressing progenitors adopt cone fates (C). *c*, In a later environment, Delta expression in progenitors results in a bias toward rod fates (R) in *d*. Inductive factors (F1–F4) change over time, so endogenous Delta signalling releases cells into the differentiation pathway in particular environments, resulting in the generation of various cell types. When progenitors misexpress Delta^{STU}, they are immune to the effects of lateral inhibition and so adopt early, GC or C, fates.

TABLE 1 Expression of cell-type markers in retinal clones

Gene	Position	No. of clones	Islet		Calbindin	
			Positive cells	No. of clones	Positive cells	
GFP	All	31	23.8 ± 1.6%	30	11.6 ± 1.3%	
Delta	Centre*	6	4.6 ± 2.3%†	8	2.1 ± 0.7%†	
Delta	Border‡	6	44.8 ± 5.0%†	8	32.0 ± 1.8%†	
Delta ^{STU}	All	32	40.7 ± 2.8%†	31	56.7 ± 3.0%†	

Antibodies against islet, a ganglion cell marker, and calbindin, a cone marker, were used to identify cell types in retinal clones. Few Delta-misexpressing cells in the centre of large (>25 cell) clones were positive for either of these markers, whereas a majority of cells at the edges of the same clones were labelled. Most of the cells misexpressing Delta^{STU} were labelled by these antibodies, regardless of clone size.

* Centre, cells more than two cell-diameters from edge of clone.

† $P < 0.001$ by Student's *t*-test.

‡ Border, cells within one cell-diameter of edge of clone. All percentages are mean ± s.e.m.

progenitor cells are induced to differentiate sequentially in a process mediated by Delta (Fig. 4); Delta expression thus controls the competence of cells to respond to inductive signals and is a key mechanism by which cell fate is regulated and histogenetic order established in the vertebrate retina. □

Methods

In situ hybridization. *In situ* hybridization was done on 10-μm paraffin sections of *Xenopus* embryos as described¹. Fluorescein-labelled antisense probe was prepared from an *X-Delta-1* cDNA (provided by A. Chitnis). For double *in situ* hybridization, a digoxigenin-labelled antisense probe of *X-Notch-1* was prepared from clone AN119 (ref. 1). Both probes were added simultaneously to prehybridized tissue sections. Probe detection was performed sequentially, with *Notch* mRNA detected first, using alkaline phosphatase-conjugated anti-digoxigenin (Boehringer Mannheim). The first signal was obtained using Fast Red substrate (Boehringer Mannheim), which produces a fluorescent red precipitate. The remaining alkaline phosphatase was inactivated by fixation and incubation for 30 min at 60 °C. Alkaline phosphatase-conjugated anti-fluorescein (Boehringer Mannheim) was added and *Delta* mRNA was detected by a blue BCIP/NBT colour reaction. Stained tissue was imaged with a Spectrasource cooled CCD camera and Adobe Photoshop software.

Gene misexpression. cDNAs for GFP(S65T), β-galactosidase, *X-Delta-1* and *X-Delta-1^{STU}* were subcloned into a pCS2+ expression vector (provided by D. Turner). The GFP(S65T) mutant was provided by R. Tsien. mRNAs transcribed *in vitro* were injected² into blastomere D.1.1 at the 16-cell stage²⁴. DNA lipofection has been described^{1,30}. *In situ* hybridization was followed by c-Myc immunostaining using the 9E10 antibody¹ for simultaneous detection of endogenous *X-Delta-1* mRNA and ectopic Delta protein. Misexpressing cells were counted and identified on the basis of their laminar position and morphology¹. Additional cell counts were done using anti-calbindin (Sigma), a cone label (W.S.C., unpublished observation), islet (ganglion cells; gift from S. Thor), rod opsin (gift from P. Hargrave), and R5 (ref. 1; Müller glia) antibodies. All methods of cell identification and counting gave similar results.

Received 10 July; accepted 11 October 1996.

- Dorsky, R. I., Rapaport, D. H. & Harris, W. A. *Neuron* **14**, 487–496 (1995).
- Coffman, C. R., Skoglund, P., Harris, W. A. & Kintner, C. R. *Cell* **73**, 659–671 (1993).
- Austin, C. P., Feldman, D. E., Ida, J. A. & Cepko, C. L. *Development* **121**, 3637–3650 (1995).
- Nye, J. S., Kopan, R. & Axel, R. *Development* **120**, 2421–2430 (1994).
- Fortini, M. E., Rebey, I., Caron, L. A. & Artavanis, T. S. *Nature* **365**, 555–557 (1993).
- Chitnis, A., Henrique, D., Lewis, J., Ish-Horowitz, D. & Kintner, C. *Nature* **375**, 761–766 (1995).
- Holt, C. E., Bertsch, T. W., Ellis, H. M. & Harris, W. A. *Neuron* **1**, 15–26 (1988).
- Belecky-Adams, T., Cook, B. & Adler, R. *Dev. Biol.* **178**, 304–315 (1996).
- La Vail, M. M., Rapaport, D. H. & Rakic, P. J. *Comp. Neurol.* **309**, 86–114 (1991).
- Turner, D. L., Snyder, E. Y. & Cepko, C. L. *Neuron* **4**, 833–845 (1990).
- Wetts, R. & Fraser, S. E. *Science* **239**, 1142–1145 (1988).
- Altshuler, D. & Cepko, C. *Development* **114**, 947–957 (1992).
- Reh, T. A. *J. Neurobiol.* **23**, 1067–1083 (1992).
- Watanabe, T. & Raff, M. C. *Development* **114**, 899–906 (1992).
- Harris, W. A. & Messersmith, S. L. *Neuron* **9**, 357–372 (1992).
- Lillien, L. *Nature* **377**, 158–162 (1995).
- Fuhrmann, S., Kirsch, M. & Hofmann, H. D. *Development* **121**, 2695–2706 (1995).
- Kelley, M. W., Turner, J. K. & Reh, T. A. *Development* **120**, 2091–2102 (1994).
- Kelley, M. W., Turner, J. K. & Reh, T. A. *Development* **121**, 3777–3785 (1995).
- Hetzler, P., Bourouis, M., Ruel, L., Carteret, C. & Simpson, P. *Development* **122**, 161–171 (1996).

21. Cabrera, C. V. *Development* **110**, 733–742 (1990).
22. Henrique, D. et al. *Nature* **375**, 787–790 (1995).
23. Nieuwkoop, P. D. & Faber, J. *Normal Table of Xenopus laevis* (North-Holland, Amsterdam, 1956).
24. Huang, S. & Moody, S. A. *J. Neurosci.* **13**, 3193–3210 (1993).
25. Cagan, R. L. & Ready, D. F. *Genes Dev.* **3**, 1099–1112 (1989).
26. Reh, T. A. & Kjaavik, I. J. *J. Neurosci.* **9**, 4179–4189 (1989).
27. Adler, R. & Hatlee, M. *Science* **243**, 391–393 (1989).
28. Reh, T. A. *J. Neurosci.* **7**, 3317–3324 (1987).
29. Waid, D. K. & McLoon, S. C. *Soc. Neurosci.* **21**, 529 (1995).
30. Holt, C. E., Garlick, N. & Cornet, E. *Neuron* **4**, 203–214 (1990).
31. Parks, A. L. et al. *Mech. Dev.* **50**, 201–206 (1995).

ACKNOWLEDGEMENTS. We thank C. Kintner and A. Chitnis for *X-Delta-1* and *X-Delta-1^{STU}* cDNAs, for communication of unpublished results and for comments on the manuscript. Research supported by the NIH. W.S.C. is supported by an NIH postdoctoral fellowship, and R.I.D. is supported by NSF and NIH predoctoral fellowships.

CORRESPONDENCE and requests for materials to W.A.H. (e-mail: harris@jeeves.ucsd.edu).

A role for Cajal–Retzius cells and *reelin* in the development of hippocampal connections

José A. Del Río*, Bernd Heimrich†, Víctor Borrell*, Eckart Förster†, Alexander Drakew†, Soledad Alcántara*, Kazunori Nakajima‡, Takaki Miyata‡§, Masaharu Ogawa§, Katsuhiko Mikoshiba‡||, Paul Derer¶, Michael Frotscher† & Eduardo Soriano*

* Department of Animal and Plant Cell Biology, Faculty of Biology, University of Barcelona, Diagonal 645, Barcelona 08028, Spain

† Institute of Anatomy I, University of Freiburg, PO Box 111, D-79001 Freiburg, Germany

‡ Molecular Neurobiology Laboratory, Tsukuba Life Science Center, The Institute of Physical and Chemical Research (RIKEN), 3-1-1 Koyadai, Tsukuba, Ibaraki 305, Japan

§ Department of Physiology, Kochi Medical School, Nankoku, Kochi 783, Japan

|| Department of Molecular Neurobiology, Institute of Medical Science, University of Tokyo, Minato-ku, Tokyo 108, Japan

¶ Developmental Neurobiology Laboratory, Pierre et Marie Curie University, 9 Quai Saint Bernard, Paris 75005, France

DURING development of the nervous system, specific recognition molecules provide the cues necessary for the formation of neural connections. In some regions, guiding cues for axonal pathfinding and target selection are provided by specific cells that exist only transiently during development, such as the floorplate or the cortical subplate^{1–4}. In the hippocampus, distinct groups of fibres innervate different layers⁵. We have tested the hypothesis that transient neurons in the hippocampus^{6,7} provide positional information for the targeting of these fibres. Here we report that ablation of Cajal–Retzius cells in organotypic slice cultures of hippocampus prevented the ingrowth of entorhinal but not of commissural afferents. Experiments inhibiting Reelin (an extracellular matrix protein expressed by Cajal–Retzius cells) and analysis of *reeler* mutant mice showed dramatic abnormalities in the development of entorhinal afferents. Thus Cajal–Retzius cells and *reelin* are essential for the formation of layer-specific hippocampal connections.

Developing hippocampal afferents innervate their specific target layer from the beginning, indicating that cues for target-layer selection are present at embryonic stages⁶. To test whether Cajal–Retzius (CR) cells⁷ are involved in the formation of hippocampal connections, we reconstituted the entorhinohippocampal pathway (EHP) using organotypic co-cultures^{8,9}. Entorhinohippocampal axons reached the hippocampus after 3 days *in*

vitro, and by days 5–7 they had densely innervated the stratum lacunosum-moleculare (SLM) in the hippocampus proper, and the outer molecular layer (OML) in the dentate gyrus (Fig. 1a). After longer incubation times the EHP became denser, but its layer specificity remained unchanged. Thus the formation of the EHP in slice cultures develops in a similar way as it does *in vivo*, indicating that positional cues for target-layer recognition by the entorhinohippocampal afferents are retained *in vitro*. Furthermore, CR cells identified by calretinin immunostaining⁷ were abundant in the SLM and OML of hippocampal slices (Fig. 1b). Anterograde tracing showed that entorhinohippocampal afferents overlapped with CR cells and formed synaptic contacts on their dendrites (Fig. 1c, d). We conclude that the temporal and spatial relationships of developing entorhinohippocampal afferents and CR cells^{6,7} are maintained in slice co-cultures.

We then attempted to ablate CR cells. Subplate neurons had been ablated in previous studies by excitotoxic lesions with glutamate-receptor agonists^{3,4}. Taking advantage of the early generation of CR cells with respect to granule cells and pyramidal neurons¹⁰ (H. Supér, A. Martínez, J.A.D.R. and E.S., unpublished results), and of their layer-specific distribution, we ablated CR cells by local application of α -amino-3-hydroxy-5-methylisoxa-

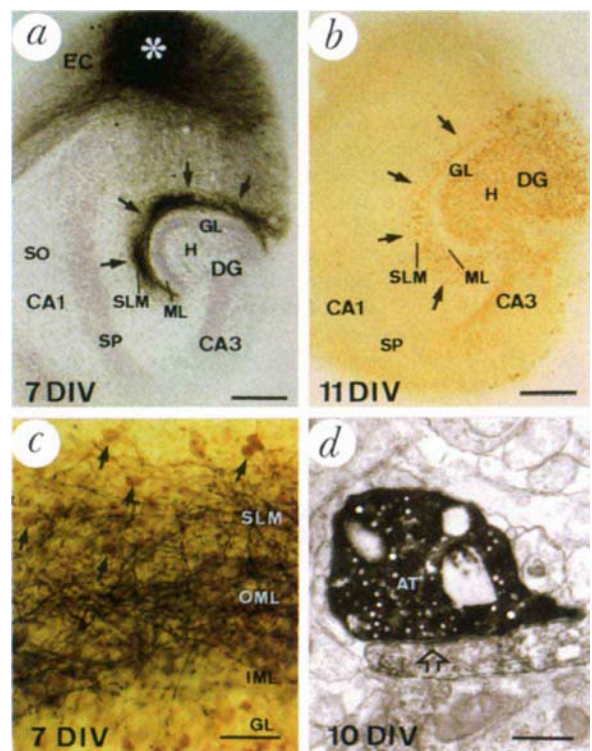


FIG. 1 Relationship between entorhinohippocampal fibres and CR cells in slice co-cultures. *a*, Laminar specificity of entorhinohippocampal fibres in a co-culture of the entorhinal cortex and hippocampus at 7 days *in vitro* (DIV). Entorhinohippocampal afferents (black, arrows) are distributed in the stratum lacunosum-moleculare (SLM) and molecular layer (ML). The injection site in the entorhinal cortex (EC) is shown by an asterisk. The section was counterstained with haematoxylin. Areas CA1 and CA3 are shown; DG, dentate gyrus; GL, granule cell layer; H, hilus; SO, stratum oriens; SP, pyramidal layer. *b*, Section from a hippocampal slice immunostained for calretinin. CR cells (arrows) are located near the hippocampal fissure, in the SLM/ML. *c*, Overlapping distribution of entorhinohippocampal axons (black) and CR cells (brown) in a hippocampal slice. Entorhinohippocampal axons impinge onto the perikarya and dendrites of Cajal–Retzius cells (arrows). IML and OML, inner and outer molecular layers. *d*, Electron micrograph showing an entorhinohippocampal axon terminal (AT) in asymmetric synaptic contact (open arrow) with a calretinin-positive process from a CR cell in the SLM. Scale bars: *a*, *b*, 100 μ m; *c*, 50 μ m; *d*, 0.5 μ m.

TABLE 1 Effects of Cajal–Retzius cell ablation and CR-50 incubation

Culture condition	Number of cultures				
(a) Ablation experiments	(–)	(+)	(++)	(+++)	Total
Control	0	0	9	20	29
Sham control	0	1	7	10	18
Lesioned	13	12	0	0	25
(b) CR-50 blocking experiments	(–)	(+)	(++)	(+++)	Total
Control	0	0	8	6	14
CR-50 (1:50)	3	7	3	0	13
CR-50 (1:100)	2	6	1	0	9
CR-50 (1:50) + <i>rl</i>	2	3	2	0	7
CR-50 (1:50) + wild-type	0	2	4	3	9
Preimmune <i>rl</i> serum (1:50)	0	0	4	3	7
Preimmune wild-type serum (1:100)	0	1	1	3	5

The EHP was classified according to the pattern of innervation in the SLM/OML: (–), virtual absence of entorhinohippocampal fibres; (+), few entorhinohippocampal fibres (<10 per 75 μ m) showing poor branching and short collaterals; (++) , moderate ingrowth of entorhinohippocampal axons (10–20 per 75 μ m) with profuse branching and long axon collaterals; (+++) , massive ingrowth of entorhinohippocampal axons (> 20 per 75 μ m) with profuse branching and collateralization. *a*, Hippocampal cocultures were lesioned and the EHP was traced with biocytin after 7 days *in vitro* (Lesioned). Controls included sham-lesioned cultures (Sham control) and unmanipulated hippocampal cultures (Control). *b*, Entorhinohippocampal cocultures were incubated with the CR-50 monoclonal antibody. After 7 days *in vitro* the EHP was traced with biocytin (CR-50). Controls included unmanipulated cultures (Control), cultures incubated with CR-50 antibody absorbed with either wild-type brain powder (CR-50 (1:50) + wild-type) or *reeler* brain powder (CR-50 (1:50) + *rl*), and co-cultures incubated with preimmune *reeler* serum and preimmune wild-type serum.

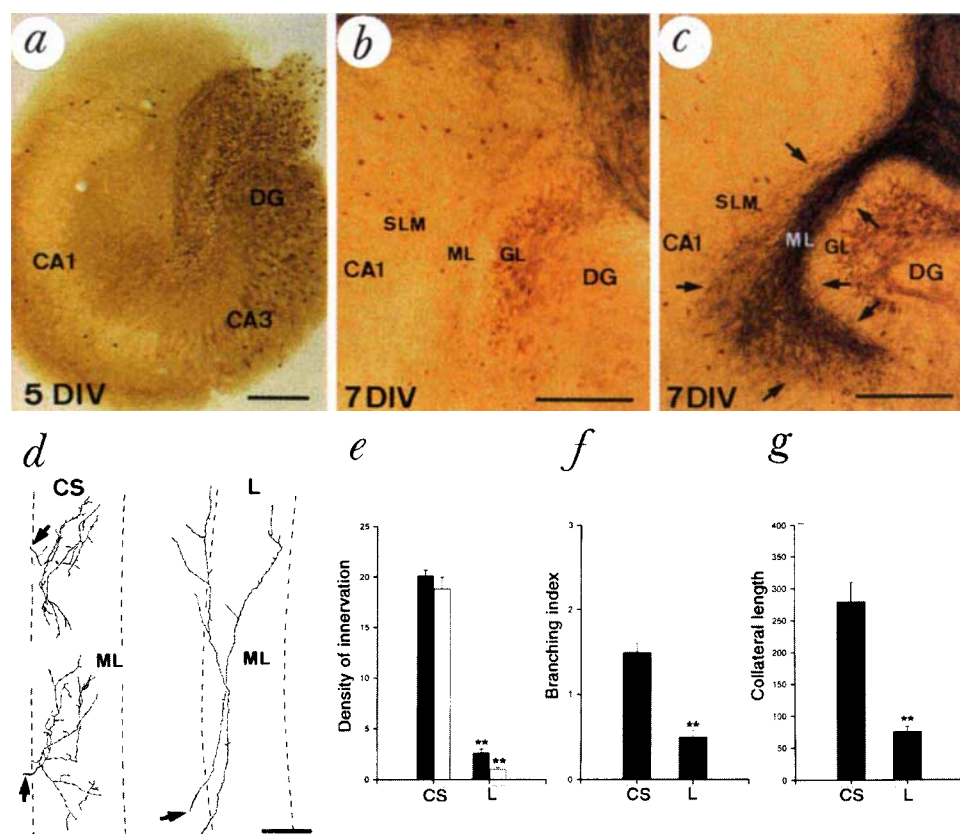


FIG. 2 Disruption of the EHP after ablation of CR cells. *a*, Hippocampal culture immunostained for calretinin after application of 6-OHDA in the SLM/OML, at 5 days *in vitro* (DIV). Note the absence of CR cells whereas the remaining layers appear normal. *b*, *c*, Sections stained for calretinin. *b*, Distribution of entorhinohippocampal afferents in a co-culture in which hippocampal CR cells had been ablated by domoic acid. Entorhinohippocampal axons are virtually absent from the SLM/ML. *c*, Distribution of entorhinohippocampal afferents (arrows) in a sham-lesioned hippocampus, showing a density and pattern of innervation indistinguishable from that of control cultures. *d*, Camera lucida drawings illustrating examples of

entorhinohippocampal axons in control sham (CS) and lesioned (L) co-cultures. Arrows point to the origin of the fibres. *e*, Density of entorhinohippocampal innervation (number of fibres per 75 μ m) in the SLM (filled bars) and OML (open bars) of control sham (CS) and lesioned (L) (mean $\pm$ s.e.m.) co-cultures. *f*, *g*, Branching index (number of axonal branching points per 100 μ m) and length of axonal side collaterals of entorhinohippocampal fibres in control sham and lesioned cocultures; significant differences between control and lesioned cultures are indicated (** P < 0.01; ANOVA, Scheffe's test). Abbreviations as in Fig. 1. Scale bars: *a*–*c*, 100 μ m; *d*, 50 μ m.

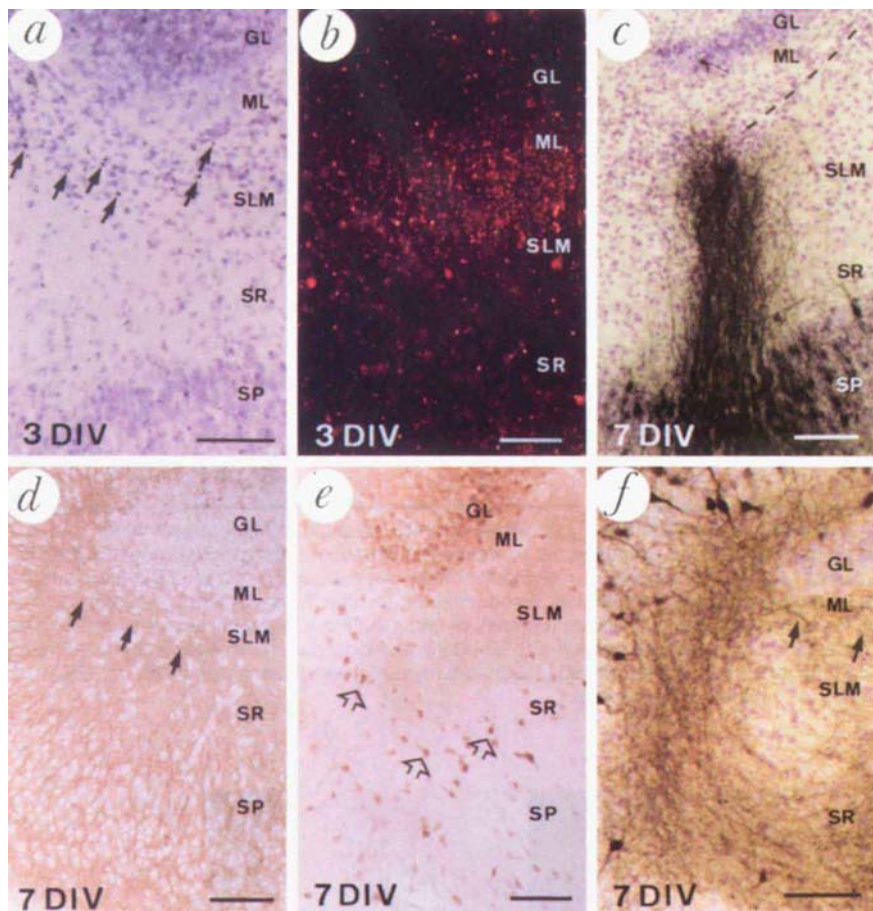


FIG. 3 *a*, Distribution of TUNEL-positive, dead cells in a hippocampal slice lesioned 3 days earlier. Most dead cells (arrows) are present within the SML/ML. Section was counterstained with haematoxylin. *b*, Confocal microscopic reconstruction illustrating the distribution of dead cells, identified by propidium iodide staining, throughout the entire depth of a hippocampal slice lesioned with 6-OHDA. *c*, Morphology of pyramidal neurons in the CA1 region in a lesioned hippocampal slice after biocytin filling in the stratum oriens. Pyramidal cells show normal dendritic arborizations reaching the hippocampal fissure. *d*, *e*, Lesioned cultures immunoreacted with MAP-2 (*d*) and calbindin (*e*) antibodies, showing that pyramidal and granule cells have a normal appearance, with dendrites that reach the SLM/ML (arrows). The calbindin-positive neurons (open arrows in *e*) residing in the stratum radiatum (SR) also survive the lesion. *f*, Distribution of associational fibres in a hippocampal slice treated with 6-OHDA. The innervation of associational axons appears normal in the SR; some fibres (arrows) were also seen in the SLM/ML. Abbreviations as in Fig. 1. Scale bars: *a*, *c*–*f*, 75 μ m; *b*, 50 μ m.

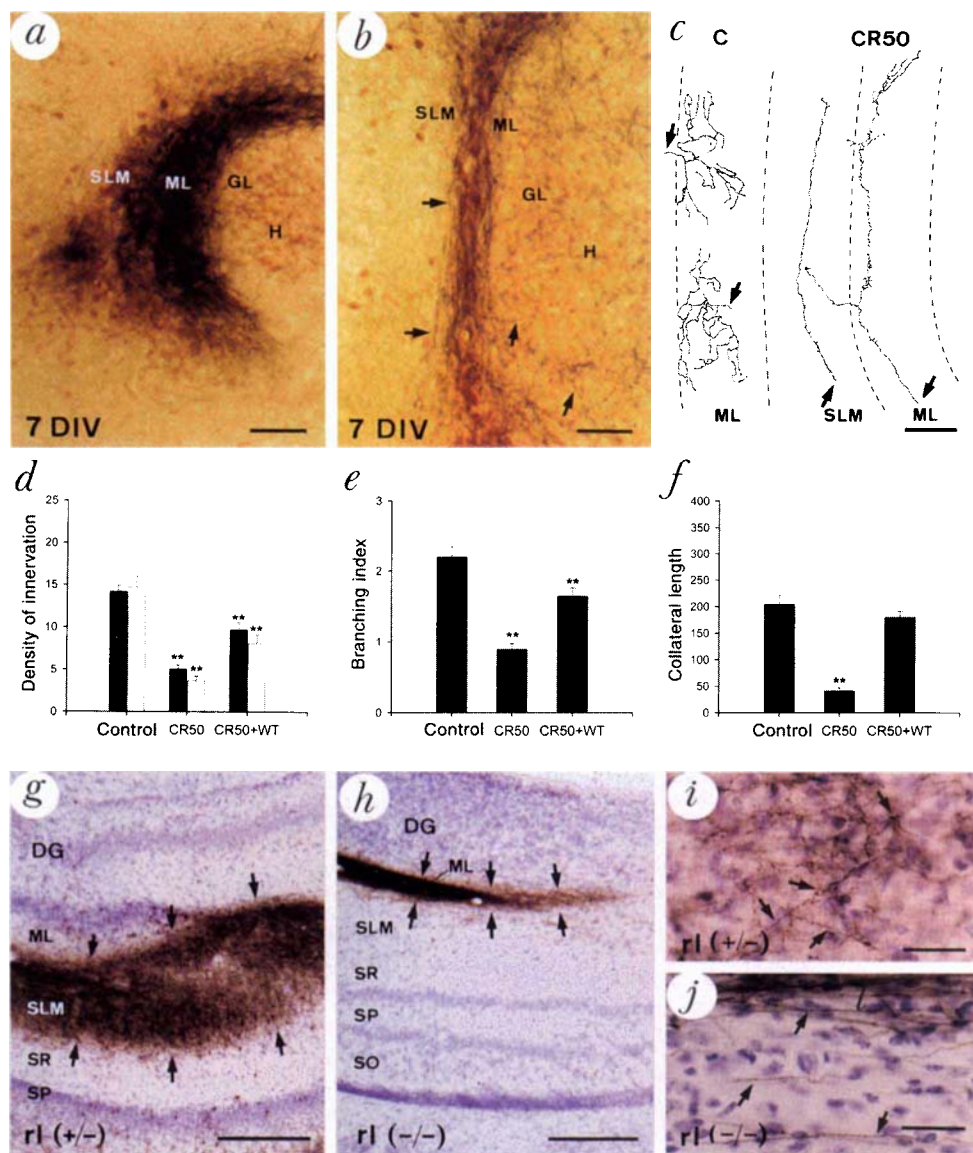
zole-4-propionic/kainate (AMPA/KA) receptor agonists, and also used 6-hydroxydopamine (6-OHDA), a toxin that causes CR cells to degenerate by a mechanism not involving AMPA/KA receptors¹¹. After 2–5 days *in vitro*, the treated slices showed a dramatic loss of CR cells identified either by calretinin immunostaining (Fig. 2*a*, *b*; –69.1%, domoic acid; –79.2%, 6-OHDA) or by pulses of 5'-bromodeoxyuridine (BrdU) administered at embryonic day 11 (–68.3%, 6-OHDA). The few CR cells that remained were shrunken and resembled dying neurons. The ablation of CR cells was confirmed by staining with propidium iodide, and terminal transferase (TdT)-mediated dUTP nick-end labelling (TUNEL) (Fig. 3*a*, *b*). In contrast, pyramidal and granule cells, filled by small injections of biocytin or immunostained for microtubule-associated protein 2 (MAP-2), had normal dendritic trees (Fig. 3*c*, *d*). Similarly, the GABA (γ -aminobutyric acid)-positive neurons of the stratum radiatum were unaffected (Fig. 3*e*). Thus our use of lesions resulted in a selective degeneration of CR cells in hippocampal cultures.

To determine whether the ablation of CR cells perturbed the formation of the EHP, lesioned hippocampal slices were co-cultured with entorhinal tissue. As in control co-cultures (Fig. 2*c*), entorhinohippocampal axons grew into the subicular region of lesioned hippocampi and ran through the alveus of areas CA1 and CA3. However, after 6–7 days *in vitro*, entorhinohippocampal axons were virtually absent in the SLM and dentate OML (Fig. 2*b*, *e*, and Table 1); moreover, the few fibres that reached these layers followed a straight course and gave off few, short axonal branches, in contrast to the profuse collateral branching and dense innervation seen in controls (Fig. 2*d*, *f*, *g*). Entorhinohippocampal fibres were also very rare in other hippocampal layers. This paucity of innervation was maintained after long incubation (15–23 days *in vitro*), suggesting that entorhinohippocampal axons were permanently unable to innervate the hippocampus (data not shown).

Similar results were obtained using either domoic acid or 6-OHDA. To determine whether these findings resulted from a general failure of axons to enter the lesioned tissue, we studied the formation of commissural/associational connections in single hippocampal cultures and in hippocampus/hippocampus co-cultures⁹. In both control and lesioned cultures, commissural/associational axons densely innervated their appropriate target layers, the stratum radiatum and oriens, and the inner molecular layer (Fig. 3*f*). In addition, in both groups some commissural/associational fibres were found in the SLM/OML (Fig. 3*f*). Thus the degeneration of CR cells does not prevent commissural/associational afferents from invading the lesioned hippocampal cultures or from growing into the lesioned SLM/OML. We conclude that the ablation of CR cells specifically affected the ingrowth and targeting of entorhinohippocampal afferents.

The gene disrupted in the *reeler* mutation is expressed by CR cells during cortical development^{12–14}. Because this gene encodes an extracellular matrix protein (Reelin), we wondered whether it might be involved in the development of the EHP. We first confirmed that the Reelin epitope recognized by the CR-50 monoclonal antibody¹⁴ was present in hippocampal CR cells *in vivo*, and found that its expression was retained in slice cultures (data not shown). Entorhinohippocampal cocultures incubated for 7–12 days *in vitro* with the blocking CR-50 antibody displayed normal cytoarchitectonics of the entorhinal cortex and hippocampus proper, whereas in the dentate gyrus, ectopic granule cells were abundant in the hilus (Fig. 4*b*). The number of entorhinohippocampal fibres crossing the subiculum towards the SLM/OML (the route of entry of the perforant pathway) was similar in control (90.37 ± 13.17 , mean $\pm$ s.e.m.) and CR-50-incubated (99.8 ± 14.38) slices, indicating that the EHP had formed. Entorhinohippocampal fibres in cultures treated with CR-50 were targeted to the SLM and OML, but they occupied a narrow

FIG. 4 Effect of the CR-50 blocking monoclonal antibody on the formation of the EHP and abnormal development in *reeler* mutant mice. **a, b**, Stained for calretinin. **a**, Development of the EHP in a co-culture incubated with preimmune *reeler* mouse serum at 7 days *in vitro* (DIV). Fibres terminate densely throughout the SLM/ML. **b**, Distribution of entorhinohippocampal axons in a co-culture incubated with the CR-50 antibody. Few entorhinohippocampal fibres (arrows) were seen in the SLM/ML. Entorhinohippocampal axons appear grouped in a narrow layer at the SLM/ML border. **c**, Camera lucida drawings showing entorhinohippocampal fibres in control (preimmune mouse serum, C) and CR-50-incubated cocultures (CR50). Arrows point to the origin of the fibres. **d**, Density of innervation (number of fibers per 75 μ m) of entorhinohippocampal fibres in the SLM (filled bars) and OML (open bars) in cultures incubated with preimmune mouse serum (Control), CR-50 antibody (1:50) absorbed with wild-type mouse brain powder (CR50 + WT), or with CR-50 antibody (CR50). **e, f**, Branching index (number of branching points per 100 μ m) and length of side axon collaterals of entorhinohippocampal fibres in the same culture groups as **d**. Data are mean $\pm$ s.e.m.; significant differences between controls and CR50+WT and CR50 cultures are indicated (** $P < 0.01$; ANOVA, Scheffe's test). **g, h**, Distribution of entorhinohippocampal afferents in the hippocampus of a heterozygous (**g**) and *reeler* (**h**) mouse at postpartum day 5. While in the heterozygous hippocampus, entorhinohippocampal fibres innervate a thick layer in the SLM/ML (arrows); in the *reeler* mutant, entorhinohippocampal axons are restricted to a thin band near the hippocampal fissure (arrows). **i, j**, Higher-magnification of entorhinohippocampal fibres at postpartum day 5. In heterozygous mice, entorhinohippocampal fibres (arrows) branch frequently, arborizing densely in the SLM/OML, whereas in *reeler* mice, fibres follow straight courses and



layer near the hippocampal fissure (Fig. 4b), in contrast to the profuse innervation seen in controls, in which entorhinohippocampal axons terminated in a thick layer throughout the SLM/OML (Fig. 4a). In addition, in CR-50-incubated co-cultures the density of innervation was reduced three- to fourfold, and single axons had significantly fewer side branches and shorter axon collaterals than did controls (Fig. 4c-f, and Table 1). Similar abnormalities were observed in cultures treated with CR-50 antibody absorbed with *reeler* tissue, but not in cultures incubated with normal mouse serum or preimmune *reeler* serum; the effects diminished notably after absorption of the CR-50 antibody with wild-type brain tissue (Fig. 4c-f; Table 1). These results suggest that Reelin is involved in the formation of the EHP by influencing axonal branching and collateral extension.

To examine whether these findings also apply to normal development *in vivo*, we traced the EHP in *reeler* mutant mice and in heterozygous and wild-type mice (controls). From postpartum day 0, the EHP had formed in *reeler* mice and entorhinohippocampal afferents were targeted mainly to the SLM/OML, indicating that

do not exhibit collateral branching (arrows). Sections are counterstained with cresyl violet. Abbreviations as in Fig. 1. Scale bars: **a, b, c, i, j**, 50 μ m; **g, h**, 100 μ m.

Reelin is not essential for either the ingrowth or the targeting of entorhinohippocampal axons. However, at all the stages examined, the zone of entorhinohippocampal termination was four- to sixfold thinner in *reeler* mutant than in control mice (Fig. 4g, h). Furthermore, up to day 5, entorhinohippocampal fibres in *reeler* mice were packed in straight axonal bundles with hardly any collateral branches, in contrast to the profuse branching seen in control mice (Fig. 4i, j). These fibre abnormalities were less apparent at days 10–12, whereas adult *reeler* mice showed branching patterns indistinguishable from those in control mice (data not shown). These results indicate that Reelin is important for the early branching and collateral extension of entorhinohippocampal axons, and that these abnormalities are recovered at later stages, probably through the action of other growth-promoting factors such as neurotrophins^{15,16}.

We have shown that ablation of CR cells largely prevents the ingrowth of entorhinohippocampal axons to the hippocampus *in vitro*, which suggests that these neurons are required for the attraction and layer-specific targeting of developing entorhino-

hippocampal afferents. This conclusion is supported by the CR-50 blocking experiments, and by the developmental study in *reeler* mice *in vivo*. Furthermore, our data suggest that the *reelin* gene is involved in axonal growth, as well as its well-known function in neuronal migration^{12–14}. Several attractive and repellent molecules, including the netrin^{1,2}, sema/collapsin^{17,18} and eph-ligand^{19,20} superfamilies, acting in conjunction with growth-permissive factors, are involved in the initial formation of neural connections. Further refinement of projections involves the action of neurotrophins^{15,16} and neuronal activity^{21,22}. The failure of entorhino-hippocampal fibres to enter their target layers in the absence of CR cells suggests that such neurons express the specific chemoattractive cues necessary for their ingrowth and terminal arborization. The observation that the EHP forms in the absence of Reelin, but with dramatic abnormalities, indicates that Reelin is not the only molecule expressed by CR cells to be involved in the formation of the EHP.

The floorplate guides commissural axons in the spinal cord², and subplate neurons in the developing neocortex are required for recognition of the target region by thalamocortical fibres^{3,4}. We have shown that another type of transient early neuron, the CR cell, is involved in target-layer selection of hippocampal afferents. □

Methods

Organotypic cultures and axonal tracing. Hippocampal cultures ($N = 32$) and entorhinohippocampal co-cultures ($n = 59$) were prepared from slices 350 μm thick taken from newborn mice (NMRI strain)^{8,9,23}. After 3–31 days *in vitro*, a crystal of biocytin was placed in the entorhinal slice. Cultures were fixed with paraformaldehyde, sectioned (40 μm) and incubated with the avidin–biotin–peroxidase complex. Sections were developed with diaminobenzidine/nickel (DAB–Ni) and counterstained with haematoxylin or immunostained for calretinin⁷ using DAB. For electron microscopy, co-cultures ($n = 16$) were postfixed with 2% osmium tetroxide and embedded in Araldite. In the electron microscope, biocytin and calretinin labelling were distinguished by the different reaction products: the DAB–Ni precipitate shows homogenous, highly electron-dense product, whereas the DAB precipitate is granular and less dense²⁴.

Lesion of hippocampal cultures. Small strips of 6% solid agar (300 μm thick) were cut on a vibratome and soaked in domoic acid (2 mM, in 0.1 PBS) or 6-OHDA (6.5 mg ml⁻¹, in 0.9% NaCl and 1 mg ml⁻¹ ascorbic acid). The agar stripes were held with fine forceps in direct contact with the SLM/OML for 5 min at the exposed surface of hippocampal slices. After several rises, hippocampal slices were cultured for 2–5 days to allow CR cells to degenerate, and then a normal entorhinal slice from a newborn mouse was added. Controls were run in parallel and included unmanipulated cultures and cultures treated in the same way, except that the agar stripes were soaked in vehicle solution (sham lesions). After 7–23 days *in vitro*, the EHP was traced with biocytin. To stain dead cells, lesioned slices ($n = 11$) were cultured for 2–3 days *in vitro* with propidium iodide (5 $\mu\text{g ml}^{-1}$)²⁵. Slices were viewed in a confocal microscope. Other cultures were processed for the TUNEL technique²⁶ ($n = 7$), and for the immunocytochemical detection of Map-2 ($n = 29$), calbindin ($n = 26$) and GABA ($n = 17$)⁷. Two pregnant mice were injected with a BrdU pulse at embryonic day 11. Pups were born and hippocampal cultures were lesioned as described. After 7 days *in vitro*, cultures ($n = 30$) were processed for the detection of BrdU²⁷. For the labelling of pyramidal neurons, lesioned hippocampal slices ($n = 14$) were injected with biocytin in the stratum oriens of CA1. Associational fibres were labelled by injecting biocytin in the pyramidal layer of CA3 ($n = 17$). For tracing commissural connections, lesioned hippocampal slices were cultured with normal hippocampal slices. After 10–15 days *in vitro*, biocytin was injected in CA3 and the hilus of the normal hippocampus ($n = 26$).

CR-50 experiments. The CR-50 monoclonal antibody¹⁴ recognizes an amino-terminal epitope of Reelin²⁸. Entorhinohippocampal cocultures from newborn mice were cultured for 7 or 12 days *in vitro* with CR-50 antibody or with mouse normal serum or *reeler* mouse preimmune serum (all used at 1:50 to 1:100). Additional controls included incubation with CR-50 antibody absorbed with brain powder from *reeler* or wild-type embryos¹⁴. After tracing the EHP with biocytin, cultures were processed for calretinin immunostaining. For CR-50 immunostaining, the brains of mice from embryonic day 18 to day 12 postpartum, and hippocampal cultures, were fixed with paraformaldehyde. Sections were incubated with CR-50 antibody (1:500 to 1:1,000)¹⁴.

Quantitative analysis. Co-cultures (7 days *in vitro*) receiving similar biocytin injections (7–9 cultures per group) were viewed with a $\times 40$ oil-immersion objective lens and a millimetric eyepiece. Axons crossing a 75- μm bar were counted in a single focusing plane (6–12 counts per culture). Selected fibres were drawn using a $\times 100$ objective lens and a camera lucida (25–43 fibres per

group). Lengths were measured using a planimeter, and the branching index (number of branching points per 100 μm) and lengths of side collaterals were calculated.

Anterograde tracing in *reeler* mice. Postnatal and adult *reeler* (r^{off} , Balb/c)²⁹, heterozygous and wild-type mice (postpartum days P0–P2, 4 $r^{(-/-)}$, 5 $r^{(-/+)}$, 7 $r^{(+/+)}$; P5, 6 $r^{(-/-)}$, 4 $r^{(-/+)}$, 6 $r^{(+/+)}$; P10–P12, 6 $r^{(-/-)}$, 5 $r^{(-/+)}$, 4 $r^{(+/+)}$; adult 7 $r^{(-/-)}$, 4 $r^{(-/+)}$, 3 $r^{(+/+)}$) were injected with biocytin in the entorhinal cortex. After perfusion with paraformaldehyde, vibratome sections (50 μm) were processed for the visualization of biocytin.

Received 20 June; accepted 11 November 1996.

- Kennedy, T. E. & Tessier-Lavigne, M. *Curr. Opin. Neurobiol.* **5**, 83–90 (1995).
- Serafini, T. et al. *Cell* **78**, 409–424 (1994).
- Soriano, E., Antonini, A., McConnell, S. K. & Shatz, C. J. *Nature* **347**, 179–181 (1990).
- Ghosh, A. & Shatz, C. J. *Science* **255**, 1441–1443 (1992).
- Amaral, D. G. & Witter, M. P. in *The Rat Nervous System* 2nd edn (ed. Paxinos, G.) 443–493 (Academic, New York, 1995).
- Supér, H. & Soriano, E. *J. Comp. Neurol.* **344**, 101–120 (1994).
- Soriano, E., Del Río, J. A., Martínez, A. & Supér, H. *J. Comp. Neurol.* **342**, 571–595 (1994).
- Li, D., Field, P. M., Starega, U., Li, Y. & Raiman, G. *Neuroscience* **52**, 799–813 (1993).
- Frotscher, M. & Heimrich, B. *Proc. Natl Acad. Sci. USA* **90**, 10400–10403 (1993).
- Soriano, E., Cobas, A. & Fairén, A. *J. Comp. Neurol.* **281**, 586–602 (1989).
- Supér, H., Martínez, A. & Soriano, E. *Dev. Brain Res.* (in press).
- D'Arcangelo, G. et al. *Nature* **374**, 719–723 (1995).
- Hirotsune, S. et al. *Nature Genet.* **10**, 77–82 (1995).
- Ogawa, M. et al. *Neuron* **14**, 899–912 (1995).
- Cabelli, R. J., Hohn, A. & Shatz, C. J. *Science* **267**, 1662–1666 (1995).
- Cohen-Cory, S. & Fraser, S. E. *Nature* **378**, 192–196 (1995).
- Kolodkin, A. L., Matthes, D. J. & Goodman, C. S. *Cell* **75**, 1389–1399 (1995).
- Messersmith, E. K. et al. *Neuron* **14**, 949–959 (1995).
- Cheng, H.-J., Nakamoto, M., Bergemann, A. D. & Flanagan, J. G. *Cell* **82**, 371–381 (1995).
- Drescher, U. et al. *Cell* **82**, 359–370 (1995).
- Goodman, C. S. & Shatz, C. J. *Cell* **72**, 77–98 (1993).
- Hermann, K. & Shatz, C. J. *Proc. Natl Acad. Sci. USA* **92**, 11244–11248 (1995).
- Stoppini, L., Buchs, P. A. & Müller, D. J. *Neurosci. Methods* **37**, 173–182 (1991).
- Gulyás, A. I. et al. *Nature* **366**, 683–687 (1993).
- Vornov, J. J., Tasker, R. C. & Coyle, J. T. *Exp. Neurol.* **114**, 11–22 (1991).
- Gavrieli, Y., Sherman, Y. & Ben-Sasson, S. A. *J. Cell Biol.* **119**, 493–501 (1992).
- Soriano, E. & Del Río, J. A. *J. Histochem. Cytochem.* **39**, 255–263 (1991).
- D'Arcangelo, G. et al. *J. Neurosci.* (in press).
- Guenet, J. L. *Mouse News Lett.* **41**, (1981).

ACKNOWLEDGEMENTS. E.S. and M.F. contributed equally to this work. We thank G. Raiman and R. Rycroft for critical reading of the manuscript. This work was supported by grants from FIS, CICYT, DICYT and CIRIT (Spain), from DFG (Germany), and from STA and MESG (Japan). V.B. was supported by the Fundació Agustí Pedró i Pons.

CORRESPONDENCE and requests for materials should be addressed to E.S. (e-mail: soriano@porthos.bio.ub.es).

Frequency detection and temporally dispersed synaptic signal association through a metabotropic glutamate receptor pathway

Andrew M. Batchelor & John Garthwaite

The Cruciform Project, University College London, 140 Tottenham Court Road, London W1P 9LN, UK

IN the classical view, a central neuron integrates incoming synaptic information by simple algebraic summation of the resultant bioelectrical signals that coincide in time. The voltage dependence of the NMDA (*N*-methyl-D-aspartate) type of ionotropic glutamate receptor endows neurons with an additional tool that allows one synaptic input to influence another, providing, again, that the two are active simultaneously¹. Here we identify a new mechanism by which non-coincident signals generated by different synaptic inputs are integrated. The device serves to regulate neuronal excitation through G-protein-coupled, metabotropic glutamate receptors (mGluRs)² in a powerful and specific manner. We show that, in cerebellar Purkinje cells, a single activation of the climbing-fibre input markedly potentiates

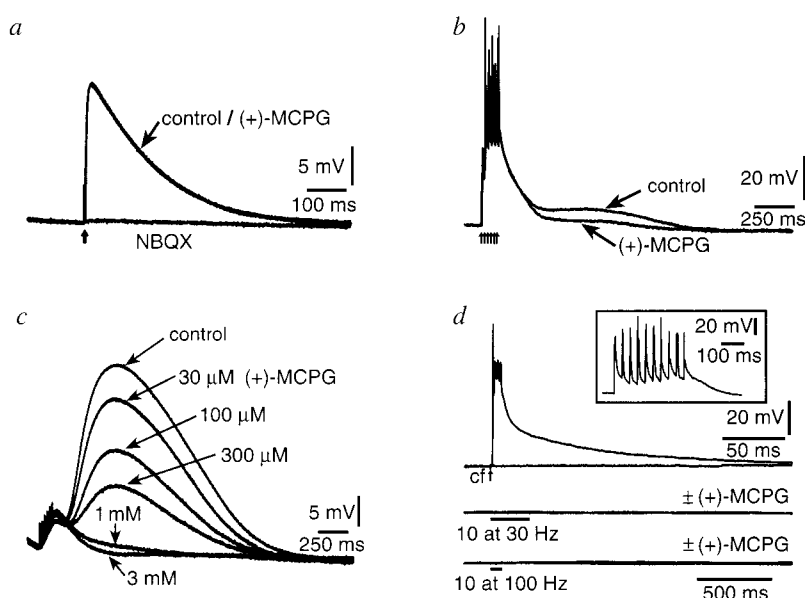
mGluR-mediated excitation at parallel-fibre synapses³. The potentiation results from a transient rise in cytosolic Ca^{2+} which is 'memorized' in such a way that it promotes excitation through mGluRs for about two minutes. A Ca^{2+} -transient is also effective if imposed up to two seconds after parallel-fibre stimulation. By allowing temporally and spatially dispersed synaptic signals to be assimilated, this mechanism adds a new element to the computational power of central neurons.

The experiments were carried out by recording intracellularly from Purkinje cells in slices of adult rat cerebellum maintained *in vitro*. Single stimuli applied to the parallel fibres gave rise to an excitatory postsynaptic potential (EPSP) that could be eliminated by NBQX (10 μM), confirming that fast neurotransmission at these synapses is mediated by the AMPA (α -amino-3-hydroxy-5-methylisoxazole-4-propionic acid) type of glutamate receptor⁴. (+)-MCPG (1 mM), an mGluR antagonist, had no effect under

these conditions (Fig. 1a). If, on the other hand, the parallel fibres were given a brief tetanic stimulation, typically 3–6 pulses at 50 Hz, a slow, delayed (+)-MCPG-sensitive depolarizing potential became apparent (Fig. 1b). This mGluR_{EPSP} could be studied in isolation after AMPA receptors had been blocked by NBQX (Fig. 1c). Relative to the first stimulus to parallel fibres, it appeared after a delay of 196 ± 15 ms, peaked at 540 ± 33 ms and decayed back to baseline after 1.5 ± 0.1 s (means $\pm$ s.e.; $n = 14$). (+)-MCPG decreased the potential in a concentration-dependent fashion (Fig. 1c), with 50% inhibition being evident at 101 ± 14 μM ($n = 3$).

The mGluR subtype responsible for the EPSP is almost certainly mGluR1 (refs 3, 5), which couples to phospholipase C and, thence, to the formation of diacylglycerol and inositol trisphosphate^{6,7}. To gain insight into the ionic mechanism underlying the mGluR_{EPSP}, its amplitude was measured as the Purkinje

FIG. 1 An mGluR_{EPSP} in cerebellar Purkinje cells associated with parallel fibre, but not climbing fibre, synaptic transmission. **a**, The EPSP evoked by low-frequency (<0.1 Hz) stimulation of parallel fibres (at vertical arrow) is blocked by the AMPA antagonist NBQX (10 μM), but not by the mGluR antagonist (+)-MCPG (1 mM). **b**, Short trains of stimuli (5–6 at 50 Hz) caused summation of the early AMPAR_{EPSP}s (leading to spike firing) and a slow, delayed, depolarizing potential which was reversibly inhibited by (+)-MCPG (1 mM). **c**, In the presence of NBQX, brief tetanic stimulation of parallel fibres evoked a response with two components. The first, which appeared during the stimulation and subsided soon afterwards, is most probably due to a raised extracellular K^+ concentration²⁵. The second component was inhibited in a concentration dependent way by (+)-MCPG. **d**, Threshold stimulation of the climbing fibre produced a typical all-or-none complex spike (top trace, one failure), which was abolished on addition of 10 μM NBQX. Tetanic stimulation (lower traces) in the presence of NBQX did not result in a detectable (+)-MCPG-sensitive response, whereas in the same cells brief tetanic stimulation of the parallel fibres always produced a (+)-MCPG-sensitive potential (not shown; $n = 5$). The inset shows that, in the absence of NBQX, the climbing fibre response is capable of following



a 30-Hz stimulation frequency. Traces in **a–c** are averages of 3–5 sweeps.

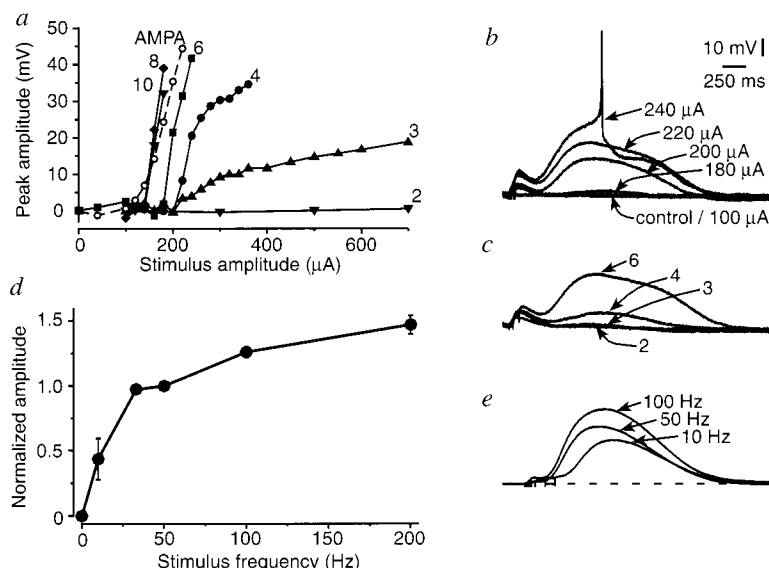


FIG. 2 Dependence of the mGluR_{EPSP} on stimulation parameters. **a**, Typical stimulus–response curves for the mGluR_{EPSP} (in the presence of NBQX) for short trains delivered at 50 Hz and containing different numbers of stimuli (as indicated). Results for the AMPAR_{EPSP} in the same cell (no NBQX) following single stimuli are shown for comparison (broken line). Four other cells behaved similarly. **b**, Sample traces of the dependence of the mGluR_{EPSP} on stimulus strength. **c**, Examples of the variation in amplitude of the mGluR_{EPSP} with the number of pulses delivered at fixed amplitude (220 μA). **d**, Frequency dependence of the mGluR_{EPSP} in response to 4 stimuli; the data have been normalized to the peak amplitude of the response at 50 Hz ($\pm$ s.e., $n = 4$); **e**, sample traces from one cell (averages of 3–4 responses).

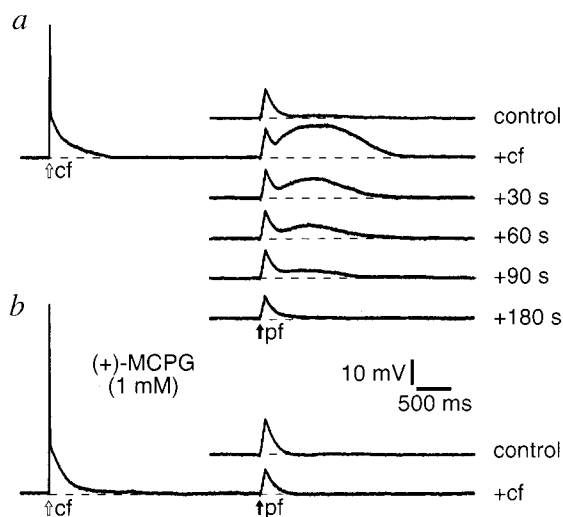


FIG. 3 Potentiation of the $mGluR_{EPSP}$ by prior stimulation of climbing fibres. **a**, Stimulation of the parallel fibres (pf) alone (4 pulses at 50 Hz) produced only a summated $AMPA_{EPSP}$ (control), whereas if a single climbing fibre (cf) response was evoked 3 s beforehand (+cf), an additional slow EPSP appeared. When only parallel fibres were subsequently stimulated (at the times indicated), the slow EPSP remained potentiated for at least 90 s. **b**, The climbing-fibre-induced potentiation of the slow EPSP was inhibited if the protocol was repeated in the presence of (+)-MCPG (1 mM). Throughout this experiment, 200 nM NBQX was present in order to suppress excessive summation of $AMPA_{EPSP}$ s.

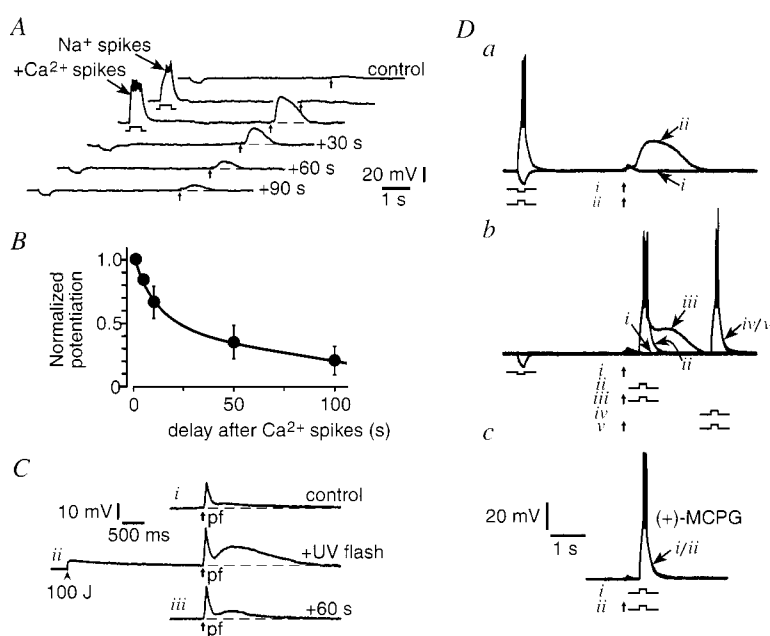
cell membrane potential was varied between -80 mV and -130 mV. The size increased linearly by 0.19 ± 0.08 mV per mV of hyperpolarization ($n = 4$), a value not significantly different from the change in size of the AMPA-receptor-mediated EPSP ($AMPA_{EPSP}$) in the same cells (0.28 ± 0.03 mV per mV; $n = 5$). When Cl^- was omitted from the electrode solution (2 M potas-

sium methylsulphate instead of 2 M KCl), the shape of the $mGluR_{EPSP}$ and its voltage dependence (0.13 ± 0.04 mV per mV; $n = 4$) were not significantly changed. These results are consistent with the principal effector mechanism being either a non-selective cation channel or an electrogenic Na^+-Ca^{2+} exchanger, as has been proposed from studies of the response of cultured Purkinje cells to an exogenous $mGluR$ agonist⁸.

Because an important determinant of the appearance, or otherwise, of the $mGluR_{EPSP}$ was the pattern of afferent-fibre activity, we characterized this parameter in more detail. One or two stimuli (at 50 Hz) were rarely effective (2 cells out of 53), regardless of the stimulus strength (Fig. 2a). With additional stimuli in a 50-Hz train, the EPSP became visible and its amplitude increased with stimulus strength (Fig. 2a, b); increasing the numbers of stimuli shifted the stimulus-response curve to the left and steepened its slope (Fig. 2a, c). At 50 Hz, 6–10 stimuli were maximally effective and the stimulus-response curve for the $mGluR_{EPSP}$ was then very similar to that of the $AMPA_{EPSP}$ observed in the same cells with single, low-frequency stimuli (Fig. 2a). This indicates that the $mGluR_{EPSP}$ does not depend on the activation of a specially large population of synapses and that $mGluRs$ can deliver as powerful a synaptic excitation to the neurons as can AMPA receptors, albeit with very different kinetics. With the numbers of stimuli in a train set at 4, the $mGluR_{EPSP}$ amplitude was steeply dependent on the stimulation frequency (Fig. 2d, e), particularly over the range (0–50 Hz) likely to be relevant to parallel fibres *in vivo*^{9,10}.

The $mGluR_{EPSP}$ appeared to be specifically associated with parallel-fibre synapses as it could not ordinarily be detected after stimulation of the other glutamatergic input to Purkinje cells, the climbing fibres (Fig. 1d). Surprisingly, however, under conditions where an $mGluR_{EPSP}$ was otherwise invisible, a single stimulation of climbing fibres revealed a large slow EPSP when parallel fibres were subsequently stimulated. This effect of climbing-fibre stimulation was observable long after the associated bioelectrical activity had disappeared, that is, for at least 90 s (Fig. 3a). When the experiment was repeated in the presence of (+)-MCPG, no slow EPSP was evident, confirming its identity as

FIG. 4 Potentiation of the parallel fibre $mGluR_{EPSP}$ by temporally dispersed Ca^{2+} spikes and the role of cytosolic Ca^{2+} . **A**, Under control conditions, parallel fibre stimulation (vertical arrow) produced a very small $mGluR_{EPSP}$ (top trace). The $mGluR_{EPSP}$ was unchanged by a prior depolarization that induced only Na^+ spike firing (second trace), but when Ca^{2+} spikes were also elicited it was greatly potentiated (third trace). The augmentation lasted for at least 90 s (timings as indicated on lower traces). Spikes are severely truncated because of filtering. **B**, Duration of the potentiating effect of prior Ca^{2+} spikes on the $mGluR_{EPSP}$. The data ($\pm$ s.e., $n = 3$) represent EPSP amplitudes normalized to their value 1 s after Ca^{2+} spiking. **C**, Photorelease of Ca^{2+} from caged Ca^{2+} potentiates the $mGluR_{EPSP}$. In control conditions (i), parallel fibre (pf) stimulation (4 stimuli at 50 Hz at vertical arrow) evoked a summated $AMPA_{EPSP}$ and a small $mGluR_{EPSP}$. Photolytic cleavage of Ca^{2+} from nitrophenyl-EGTA/ Ca^{2+} complex (100 J discharge at arrowhead) caused a large enhancement of the $mGluR_{EPSP}$ (ii). The augmentation was still visible when parallel fibres were stimulated 60 s afterwards (iii). The increased size of the $AMPA_{EPSP}$ evident after flash photolysis in this experiment was not a consistent finding ($n = 3$). **D**, Potentiation of the $mGluR_{EPSP}$ by delayed Ca^{2+} spiking. **a**, Under control conditions (i), the stimulus intensity delivered to the parallel fibres (4 pulses at 50 Hz; vertical arrows) was subthreshold for the $mGluR_{EPSP}$, whereas a prior depolarization sufficient to cause Na^+ and Ca^{2+} spikes caused a large EPSP when parallel fibres were stimulated 3 s later (ii). **b**, A depolarizing pulse (whose effect on its own is shown in traces ii and iv), when imposed 0.5 s after a normally subthreshold parallel fibre stimulation (i) caused the appearance of a depolarizing wave (iii) similar in amplitude and time course to the corresponding part of the potentiated



$mGluR_{EPSP}$ (see a), whereas it was ineffective if applied 2.5 s after synaptic stimulation (trace v). **c**, This slow depolarizing wave was abolished by (+)-MCPG (1 mM).

the mGluR_{EPSP} (Fig. 3b).

A series of tests was made to understand how the climbing-fibre input influences the parallel-fibre mGluR_{EPSP}. Presynaptic activity in climbing fibres *per se* was not sufficient because no potentiation was observed when the climbing-fibre-induced EPSP was eliminated by NBQX (10 μ M) (results not shown). If, under these conditions, Purkinje cells were directly depolarized (for 400 ms) to a level that generated only Na⁺ spikes, potentiation was still not produced. When the depolarization was increased so as to elicit firing of both Na⁺- and Ca²⁺-dependent spikes, however, a large slow EPSP was observed on subsequent parallel-fibre stimulation (Fig. 4A). The duration of the augmentation brought about in this (~2 min) was indistinguishable from that observed following climbing-fibre stimulation (Fig. 4B). This result indicates that Ca²⁺ spike firing is critical for depolarization-induced potentiation and, because firing of Ca²⁺ spikes is a physiological response of Purkinje cells to climbing-fibre activity^{11,12}, it can be concluded that this event also underlies the climbing-fibre-induced potentiation of the mGluR_{EPSP}.

One of the consequences of Ca²⁺ spike firing is a transient rise in cytosolic Ca²⁺ concentration^{11,12}. To investigate the significance of this step, two tests were done. In one, Ca²⁺ buffer was introduced into Purkinje cells by including EGTA (200 mM) in the recording electrode; after 20–30 min, the mGluR_{EPSP} (with or without preceding Ca²⁺ spikes) was abolished, whereas the AMPAR_{EPSP} persisted ($n = 3$; results not shown). In the other, Ca²⁺ was liberated in Purkinje cells by flash photolysis of caged Ca²⁺ in the form of a nitrophenyl-EGTA/Ca²⁺ complex¹³; this procedure was able to mimic the potentiating effect of Ca²⁺ spikes on the mGluR_{EPSP} (Fig. 4C). These findings signify that the mGluR_{EPSP} requires Ca²⁺ and that it is the rise in cytosolic Ca²⁺ generated by Ca²⁺ spike firing that is responsible for the climbing-fibre-induced potentiation of the mGluR_{EPSP}.

As well as being effective in advance of parallel-fibre synaptic activity, Ca²⁺ spiking could also act afterwards, even at times when conventional neurotransmission at this synapse is complete and the glutamate concentration in the synaptic cleft is assumed to be back to its resting level. The window for this delayed action of Ca²⁺ spikes was, however, restricted to the normal timespan of the mGluR_{EPSP}, that is, 1–2 s after parallel-fibre stimulation (Fig. 4D). Thus, the window for modulation of the mGluR_{EPSP} by Ca²⁺ spikes is asymmetrical with respect to time.

We have identified a new mechanism for synaptic and inter-synaptic signal assimilation that endows neurons with previously unrecognized computational abilities. Because mGluR1 is widely expressed through the central nervous system⁶, the phenomenon is likely to be of significance outside the cerebellum; its relevance may conceivably extend to the other phospholipase C-coupled mGluR, mGluR5 (ref. 14), as well as to analogous G-protein-linked receptors for other neurotransmitters, such as acetylcholine and monoamines. Mechanistically, the device relies on non-bioelectrical traces of the two inputs persisting over time. For one of the inputs, climbing fibres in this case, the relevant step is the firing of Ca²⁺ spikes that propagate through the dendrites, causing a rise in cytosolic Ca²⁺. It is unlikely that cytosolic Ca²⁺ itself is the trace, however, because the dendritic Ca²⁺ response to a Ca²⁺ spike lasts only for about 1 s (refs 15, 16), whereas its effect on a subsequent mGluR_{EPSP} lasts for two orders of magnitude longer. How the Ca²⁺ transient is memorized is unknown. Among existing mechanisms, one possibility is that Ca²⁺ causes an enduring translocation of protein kinase C from the cytosol into the membrane, which would position the enzyme for activation by diacylglycerol generated by mGluR1-coupled phospholipase C^{17,18}. Protein kinase C might then phosphorylate and activate membrane proteins, generating the EPSP. The trace of the other input is more ephemeral because the time window permitted for delayed Ca²⁺ spikes to engage the mGluR_{EPSP} is shorter (2 s); it could be a biochemical signal molecule (for example, diacylglycerol) associated with the mGluR transduction pathway. For the EPSP to be elicited homosynaptically, the Ca²⁺ signal may

initially derive from intracellular stores in response to inositol trisphosphate⁷ and be subsequently boosted by entry through local voltage-sensitive Ca²⁺ channels¹⁹.

In the cerebellum, the function of the powerful one-to-one climbing-fibre input to Purkinje cells has so far remained elusive, although it has long been regarded on theoretical grounds as a likely modifier of incoming signals from parallel fibres^{20,21}. The identification of a substrate for such an interaction is likely to help understand, *inter alia*, why activation of both sets of synapses is necessary for the induction of long-term depression of parallel fibre–Purkinje cell synaptic transmission²², a well known example of synaptic plasticity that is thought to be relevant to motor learning behaviour and is known to require the climbing-fibre-mediated increase in cytosolic Ca²⁺ and the presence of mGluR1^{5,23}. □

Methods

Purkinje cells in biplanar²⁴ or sagittal slices (350–400 μ m thick) of adult (>4 weeks old) rat cerebella were impaled with microelectrodes (60–110 M Ω , filled with 3M KCl unless otherwise stated). Cells were hyperpolarized to an apparent membrane potential of –80 mV using current injection (–0.2 to –1.23 nA) except where otherwise stated. Data was collected by an Axoclamp-2B amplifier (Axon Instruments) which was interfaced (TL-1, Axon) to a 80486-based personal computer running Axotape software (Axon). The slices were held submerged at 30 °C in a chamber (~200 μ l) that was perfused at a rate of 1 ml min^{–1} with artificial cerebrospinal fluid (aCSF) containing (in mM): NaCl 120; KCl 2; CaCl₂ 4; NaHCO₃ 26; MgSO₄ 1.19; KH₂PO₄ 1.18; D-glucose 11, equilibrated with 5% CO₂ in O₂. (–)Bicuculline methochloride (50 μ M; Tocris-Cookson, Bristol, UK), D-2-amino-5-phosphopentanoate (30 μ M; Tocris-Cookson) and CsCl (2 mM) were also added to block GABA_A receptors, NMDA receptors and the hyperpolarization-activated non-selective cation conductance, respectively. These compounds did not affect the mGluR_{EPSP}. When used, the AMPA antagonist 6-nitro-7-sulphamoylbenzo[f]quinoxaline-2,3-dione (NBQX; NovoNordisk, Denmark) and (+)- α -methyl-4-carboxyphenylglycine ((+)-MCPG; Wellcome Research) were added to the perfusate. Parallel fibres and climbing fibres were stimulated with constant-current pulses of 300 μ s duration; unless otherwise stated, stimuli were repeated every 10 s and 30 s, respectively, for the AMPAR_{EPSP} and the mGluR_{EPSP}. For flash photolysis, electrode tips were filled with nitrophenyl-EGTA (50 mM; Molecular Probes) pre-loaded with 25 mM Ca²⁺. At least 20 min after impalement of Purkinje cells, Ca²⁺ was liberated by the discharge of a xenon flashlamp (XF-10; Hi-Tech Scientific, Salisbury, UK) equipped with a UV bandpass filter (centre, 320 nm). Control experiments without nitrophenyl-EGTA showed that the 100 J flash elicited a similar artefact (Fig. 4C, trace ii) but had no effect on the mGluR_{EPSP} under conditions where depolarization-induced Ca²⁺ spikes caused the usual large enhancement.

Received 14 August; accepted 10 October 1996.

- Bliss, T. V. P. & Collingridge, G. L. *Nature* **361**, 31–39 (1993).
- Tanabe, Y., Masu, M., Ishii, T., Shigemoto, R. & Nakanishi, S. *Neuron* **8**, 169–179 (1992).
- Batchelor, A. M., Madge, D. J. & Garthwaite, J. *Neuroscience* **63**, 911–915 (1994).
- Liano, I., Marty, A., Armstrong, C. M. & Konnerth, A. *J. Physiol. (Lond.)* **434**, 183–213 (1991).
- Conquet, F. et al. *Nature* **372**, 237–243 (1994).
- Masu, M., Tanabe, Y., Tsuchida, K., Shigemoto, R. & Nakanishi, S. *Nature* **349**, 760–765 (1991).
- Berridge, M. J. *Nature* **361**, 315–325 (1993).
- Staub, C., Vranesic, I. & Knöpfel, T. *Eur. J. Neurosci.* **4**, 832–839 (1992).
- Merrill, E. G., Wall, P. D. & Yaksh, T. L. *J. Physiol. (Lond.)* **284**, 127–145 (1978).
- Eccles, J. C., Ito, M. & Szentágothai, J. *The Cerebellum as a Neuronal Machine* (Springer, Berlin, 1967).
- Llinás, R. & Sugimori, M. *J. Physiol. (Lond.)* **305**, 171–195 (1980).
- Ross, W. N. & Werman, R. *J. Physiol. (Lond.)* **369**, 319–336 (1987).
- Ellis-Davies, G. C. R. & Kaplan, J. H. *Proc. Natl Acad. Sci. USA* **91**, 187–191 (1994).
- Abe, T. et al. *J. Biol. Chem.* **267**, 13361–13368 (1992).
- Miyakawa, H., Lev-Ram, V., Lasser-Ross, N. & Ross, W. N. *J. Neurophysiol.* **68**, 1178–1189 (1992).
- Sugimori, M. & Llinás, R. R. *Proc. Natl Acad. Sci. USA* **87**, 5084–5088 (1990).
- Huang, K.-P. *Trends Neurosci.* **12**, 425–432 (1989).
- Marsh, S. J., Trouslard, J., Leaney, J. L. & Brown, D. A. *Neuron* **15**, 729–737 (1995).
- Eilers, J., Augustine, G. J. & Konnerth, A. *Nature* **373**, 155–158 (1995).
- Albus, J. S. *Math. Biosci.* **10**, 25–61 (1971).
- Marr, D. A. *J. Physiol. (Lond.)* **202**, 437–470 (1969).
- Ito, M. *Annu. Rev. Neurosci.* **12**, 85–102 (1989).
- Sakurai, M. *Proc. Natl Acad. Sci. USA* **87**, 3383–3385 (1990).
- Garthwaite, J. & Batchelor, A. M. *J. Neurosci. Meth.* **64**, 189–197 (1996).
- Batchelor, A. M. & Garthwaite, J. *J. Neuropharmacol.* **32**, 11–20 (1993).

ACKNOWLEDGEMENTS. This work was supported by a grant from the International Human Frontiers Science Program. We thank D. J. Madge (The Cruciform Project) and T. Honoré (NovoNordisk) for supplying (+)-MCPG and NBQX, respectively.

CORRESPONDENCE and requests for materials should be addressed to J.G. (e-mail: john.garthwaite@ucl.ac.uk).

Potent antitumour activity of a new class of tumour-specific killer cells

Si-Yi Chen*, An-Gang Yang*, Ji-Dai Chen*, Timothy Kute†, C. Richter King‡, John Collier§, Yanping Cong*, Changping Yao* & Xue F. Huang||

Departments of * Cancer Biology, † Pathology, and || Comparative Medicine, Comprehensive Cancer Center, Bowman Gray School of Medicine, Wake Forest University, Winston-Salem, North Carolina 27157, USA

‡ Department of Biochemistry, Lombardi Cancer Center, Georgetown University, Washington DC 20007, USA

§ Department of Microbiology and Molecular Genetics, Harvard Medical School, Boston, Massachusetts 02115, USA

Two approaches to the antibody-directed targeting of toxic or cytolytic activity and augmentation of cellular immune responses have been explored for tumour immunotherapy, but so far success has been limited¹⁻³. Obstacles facing immunotherapy are the limited accessibility of antibodies or antibody conjugates to solid tumours and the difficulty in obtaining tumour-specific cytotoxic lymphocytes⁴⁻⁷. Here we generate a new class of tumour-specific killer cells by genetically modifying lymphocytes to produce and secrete a targeted toxin against an oncoprotein overexpressed on breast and other tumour cells. The transduced lymphocytes were shown to have potent and selective cytotoxicity to tumours in culture and nude mouse models. The potent *in vivo* antitumour activity is probably a result of the migration of the lymphocytes to tumours as a targeted toxin carrier, and production and accumulation of the targeted toxins inside tumours as a producer. Our approach, which has features of both antibody-directed and cell-mediated immunotherapy, may have application in a gene therapy context.

Toxin molecules, such as *Pseudomonas* exotoxin A (PEA), kill a cell by inactivating the elongation factor-2 (EF-2) in the cytosol^{2,8}, suggesting that mammalian cells can be genetically modified to produce and secrete targeted toxins by using a signal sequence to lead newly synthesized toxins into the lumen of the endoplasmic reticulum (ER) co-translationally⁹. The toxin-expressing cells remain viable because the interaction of the luminal toxins with EF-2 is blocked by the membrane bilayer of the ER and secretory vesicles. The secreted toxins selectively destroy target cells after being internalized and released into the cytosol^{12,8}, but cannot kill the toxin-expressing cells lacking the target antigen.

A single-chain antibody (sFv) gene comprising the linked heavy and light chain variable regions derived from a monoclonal antibody e23 with high-affinity binding to the extracellular domain of HER-2 overexpressed on various human tumour^{10,11} was fused in-frame with a truncated PEA gene (PE40) encoding domains II (translocation) and III (catalytic)^{2,8} (Fig. 1A). To investigate whether mammalian cells can produce targeted toxins, COS cells were transfected with pCMV-e23sFv-PE40 and analysed by immunofluorescent staining. Strong positive staining in the cytoplasm, especially in the perinuclear Golgi region, of the transfected cells was observed when incubated with an anti-e23sFv¹¹ or anti-PEA antibody (Fig. 1B), resembling a typical secretory protein pattern¹². The expression of the fusion toxin was further demonstrated by radiolabelling and immunoprecipitation analysis (data not shown).

HER2-specific killer cells were generated by transducing pCMV-e23sFv-PE40 into human T lymphocytes (MOLT-4), which express undetectable levels of HER-2 (refs 13-15, and data not shown), followed by G418 selection¹². Pulse-chase radiolabelling showed that the fusion toxins were produced and secreted from transduced cells (Fig. 2a). Genomic polymerase chain reaction (PCR) analysis showed that the toxin gene was

specifically amplified from the transduced cells. A significant toxin enzymatic activity was detected in the MOLT-e23sFv-PE40 medium, but the activity was only background level in the medium of control cells (Fig. 2b). By comparison with the standard curve of purified PEA activity, over 0.5 µg PEA was produced from MOLT-e23sFv-PE40 at 24 h per 1×10^6 cells ml⁻¹. Biological features of MOLT-e23sFv-PE40, including cell viability, proliferation rate and protein synthesis rates were similar with untransduced lymphocytes.

To determine the selective cytotoxicity, cell lines expressing different HER-2 levels were co-cultivated with MOLT-e23sFv-PE40. Cell killing was significant in tumour cells overexpressing HER-2 (refs 16, 17) co-cultivated with MOLT-e23sFv-PE40, whereas only a small proportion of the cells expressing low levels of HER-2 (ref. 18) were killed in the co-cultures (Fig. 3a). The MOLT control had no significant effects on tumour cells expressing either high or low levels of HER-2. The specificity of the cytotoxicity was confirmed by competitive inhibition using the parental e23 antibody (Fig. 3b). A tumour xenograft nude mouse model was used to evaluate the *in vivo* antitumour activity. The administration of MOLT-e23sFv-PE40 significantly inhibited tumour growth (not shown) and prolonged animal survival (Fig. 4a, a).

To evaluate the potential application *in vivo*, human LAK cells derived from peripheral blood mononuclear cells (PBMCs)¹⁹ were transduced by recombinant retroviral vectors²⁰. SKOV3 xenograft mice were given either transduced or mock-transduced LAK weekly for two weeks. These mice were also injected with a modified e23sFv-PE40 immunotoxin with a 6-10-fold increase in cytotoxicity over e23sFv-PE40 (refs 11, 15), using an amount (0.15 µg) comparable to the estimated e23sFv/PE40 produced by the transduced cells (0.15 µg per 10^6 cells over 24 h) daily for two weeks. As shown in Fig. 4a, b, tumour growth was effectively inhibited by the transduced cells but not by direct administration of the same amount of immunotoxin, in agreement with previous studies^{11,15}.

We next looked at the *in vivo* distribution of LAK cells: toxin-

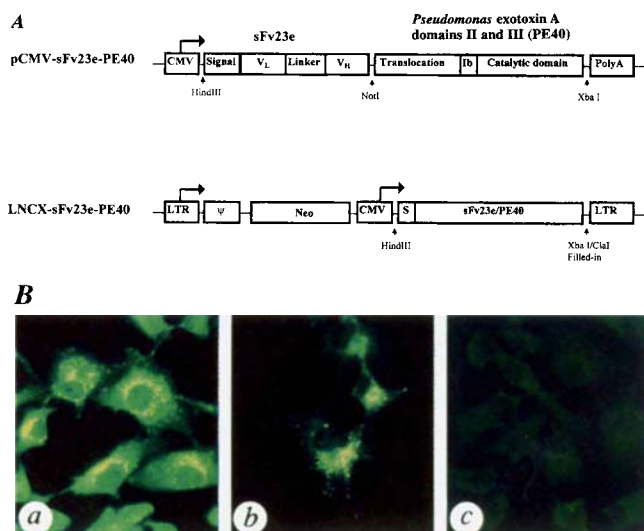


FIG. 1 Expression of recombinant anti-HER-2 fusion toxins in mammalian cells. A, The e23sFv gene¹¹ with a leader signal sequence and the PE40 (amino acids 253 to 613 of PEA) gene were cloned into the pRc/CMV vector with a neo marker (Invitrogen), or into a murine retroviral vector LNCX²⁷. These constructs were confirmed by DNA sequencing (cloning procedure available from S.-Y.E. upon request). B, Immunofluorescent staining showed that transfected COS cells with normal morphology expressed the fusion toxins. a, b, COS transfected with pCMV-e23sFv-PE40, anti-PEA (Gibco BRL) (a), or anti-e23sFv¹¹ (b); c, COS with control vector, anti-e23sFv and -PEA mixture. Magnification, $\times 800$.

expressing LAK cells were detected in tumour tissues 1 h after injection and were still evident one week later by immunofluorescent staining (Fig. 4B, *a-d*). Only low toxin enzymatic activity was found in the blood, but activity was increased in the tumours after cell injection (Fig. 4B, *e*). There was no histological abnormality in liver, lung, kidney or heart, although toxin-expressing cells were detected in these tissues (not shown). In contrast, in mice given immunotoxin directly, significant toxin activity was only transiently detected in blood and tumours after injection (Fig. 4B, *e*), suggesting that e23sFv-PE40, which is continuously produced by the transduced cells inside tumours, accumulated locally to exert its antitumour effects.

Our results show that *in vivo* production of targeted toxins by mammalian cells is feasible and has potential therapeutic application. In one application, lymphocytes transduced to secrete targeted toxins combine antibody specificity, toxin potency, and the effector-cell properties of lymphocytes. The production and accumulation of targeted toxins inside tumours by migrated lymphocytes exert more selective and potent antitumour activity

compared with administration of immunotoxin directly. Micro-metastases can be selectively destroyed by our transduced lymphocytes when reinfused back into patients, having an immune surveillance function. Our approach could be applied locally, for example to transfect neurons to make them secrete toxins to kill malignant cells (manuscript in preparation). Given that almost any antigen can be targeted on the cell surface, this approach should be applicable to treatment of cancer, autoimmune diseases and virus infection²¹. □

Methods

Immunofluorescent staining, radiolabelling and immunoprecipitation. For fluorescent staining, COS cells on coverslips were transfected with DNA, and stained as described¹². Frozen sections of tissues were stained with the anti-PEA antibody. Transduced cells were incubated with ³⁵S-cysteine for various times, and culture medium and cell lysates were precipitated with the antibodies. Samples were analysed by electrophoresis on SDS-polyacrylamide gels²² and visualized by a Phosphorimager (Molecular Dynamics).

Retroviral vectors and gene transfer. PBMCs from healthy individuals were separated on a Histopaque density gradient, and non-adherent LAK cells were

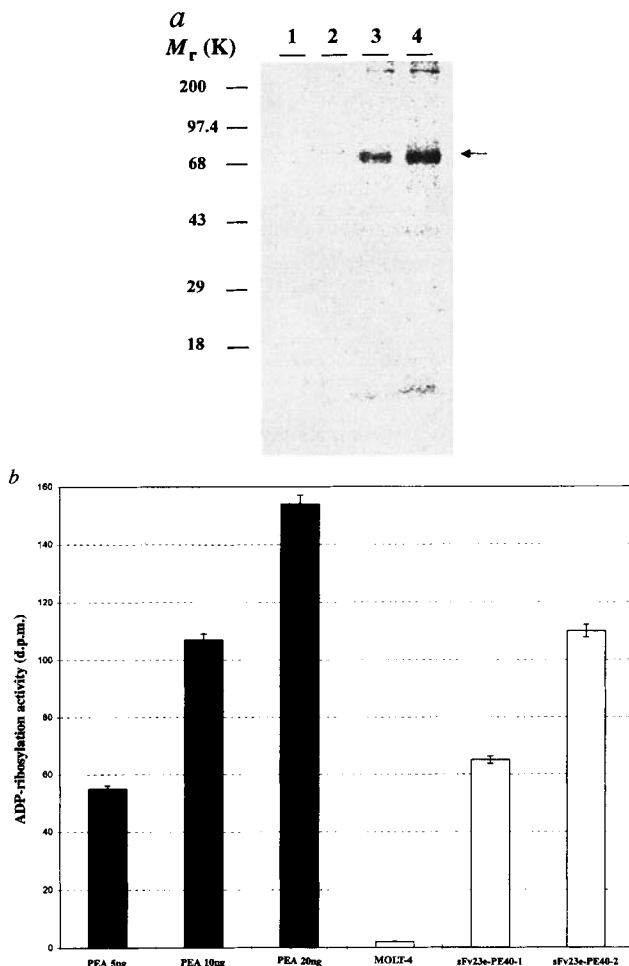


FIG. 2 Generation and analysis of tumour-specific killer cells. *a*, Secretion of fusion toxins from transformed lymphocytes. Lymphocytes (1×10^7) were pulse-radiolabelled with ³⁵S-cysteine (200 μ Ci) for 30 min and chased for various times. The culture medium was precipitated with a mixture of anti-e23sFv and anti-PEA and analysed by SDS-PAGE. Lanes: 1, MOLT control cells, chased for 4 h; 2–4, MOLT-e23sFv-PE40; chase was for 0 min, 1 h, and 4 h, respectively. *b*, ADP-ribosylation activity of secreted fusion toxins. The culture medium of MOLT-e23sFv-PE40 without (e23sFv-PE40-1) or with (e23sFv-PE40-2) concentration by Amicon filtration was assayed for ADP-ribosylation activity. Purified PEA proteins were used as positive controls. The mean scintillation counts of the samples are calculated from duplicate determinations after subtracting the background.

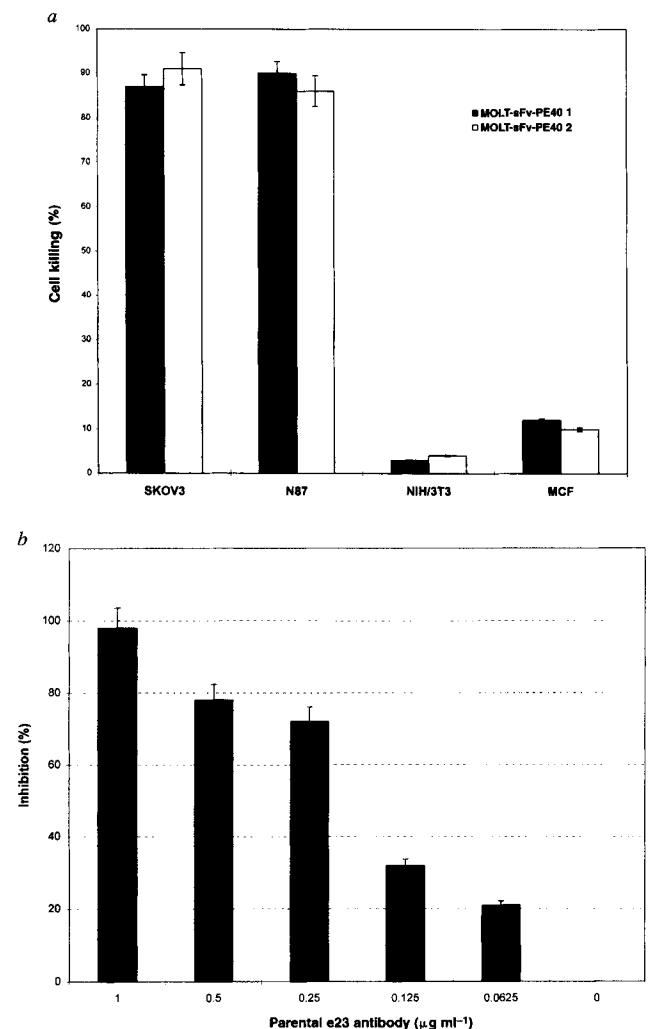


FIG. 3 Selective cytotoxicity to tumour cells *in vitro*. Tumour cells (SKOV-3, N-87) overexpressing HER-2, and control cells (MCF-7 and NIH3T3) expressing little or undetectable HER-2 (refs 16–18) on 96-well plates (1×10^5 per ml) were cultured with MOLT-e23sFv-PE40 (clones 1 and 2) or MOLT control cells (5×10^5 per ml) for 62 h. *a*, Percentage of selective cell killing. Different amounts of parental e23 antibody were added into co-cultures of MOLT-e23sFv-PE40 (clone 1) and SKOV3. *b*, Percentage of inhibited cell killing.

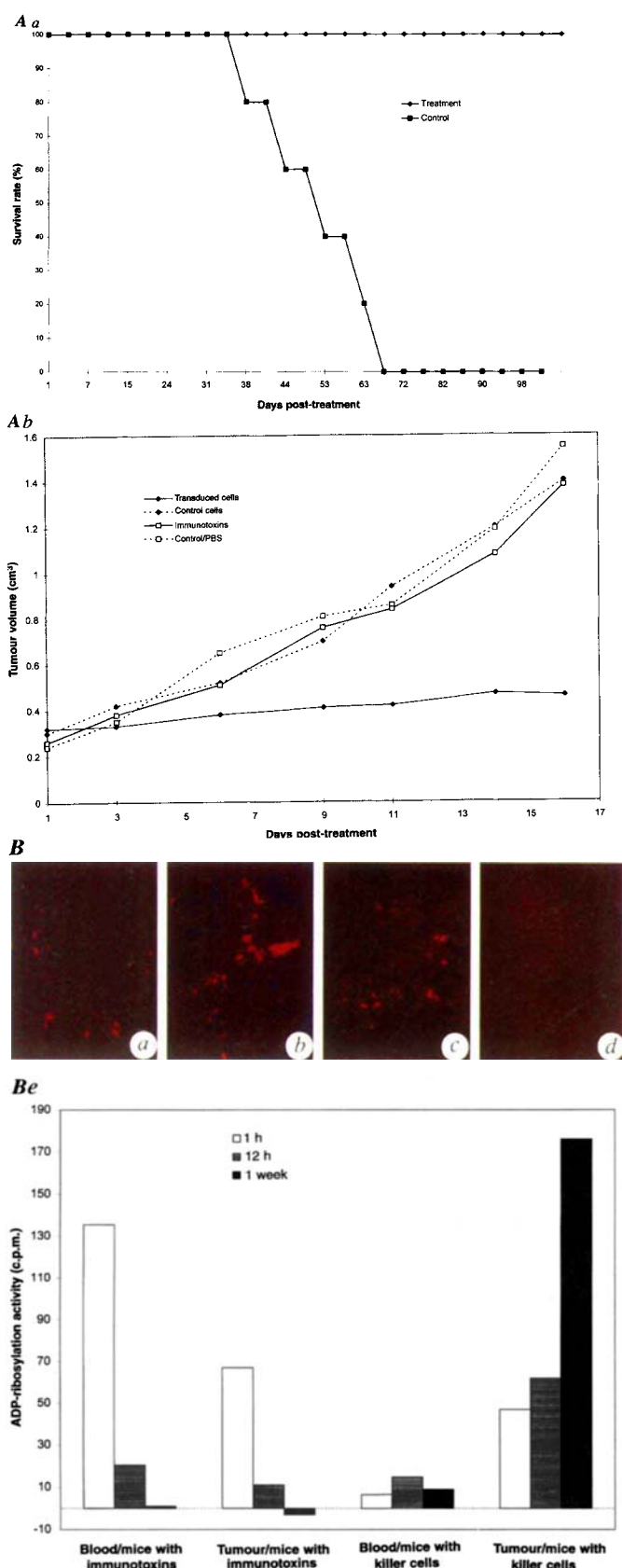


FIG. 4 Anti-tumour activity in nude mouse models. **A**, Inhibition of tumour growth and prolonged animal survival. **a**, N87 xenograft mice were given MOLT-e23sFv-PE40 (7 mice) or MOLT control cells (5×10^7) on day 1, followed by weekly injection for 6 weeks. **b**, SKOV3 xenograft mice were divided into four groups: two groups (6 mice each) were administered transduced or mock-transduced LAK cells (1×10^6) at day 1, followed by weekly injection once, and two groups were administered with 0.15 μ g of the purified immunotoxins (6 mice) or PBS (5 mice) daily for two weeks. Survival rates and tumour growth curves are shown. **B**, Toxin enzymatic activities in

generated by incubating in a culture medium (RPMI-1640/10% FCS supplemented with rIL-2 ($1,000 \text{ IU ml}^{-1}$) (Proleukin, Chiron) and PHA ($5 \mu\text{g ml}^{-1}$) for 24 to 48 h (ref. 19). The recombinant retroviral vector (Fig. 1A) was transfected into the PA317 amphotropic packaging cell line. Forty-eight hours later, the transduced cells were co-cultivated with human LAK cells in the culture medium containing protamine sulphate ($5 \mu\text{g ml}^{-1}$) for 24 h (ref. 20). LAK cells were collected and selected for one week in $300 \mu\text{g ml}^{-1}$ G418 (Gibco BRL), followed by expansion in the culture medium. Genomic PCR and immunofluorescent staining were used to detect the gene transduction and toxin proteins in the transduced lymphocytes.

ADP ribosylation assay and cytotoxicity assay in vitro. Culture media of transduced cells, or blood and tumour homogenates, were assayed for ADP-ribosylation activity as described^{23,24}. For *in vitro* cytotoxicity assays, tumour cells on 96-well plates were co-cultured with lymphocytes for 62 h. The percentage of cell killing was determined as follows: $100\% - \text{numbers of viable cells in the killer and target cell co-culture/numbers of viable cells in the control and target cell co-culture} \times 100\%$. The percentage of cell killing is presented after subtracting background cell death. For competitive inhibition assays, the percentage of cell-killing inhibition was determined: $100\% - \text{percentages of cell killing in the co-culture with the e23 antibody/percentage of cell killing in the co-culture without e23} \times 100\%$.

Antitumour assay in tumour xenograft nude mouse models. Tumour cells (5×10^6) were injected s.c. into the flanks of athymic/Nu mice (4 to 6 weeks old) (Goodwin Institute for Cancer Research). Two to three weeks later, the tumour xenografts were implanted into the flanks of athymic mice, and grown for one to two weeks. Mice with relatively uniformly grown tumour xenografts were randomized to treatment and control groups. Tumour diameters were caliper-measured twice a week, and tumour volume was calculated from the formula: $\text{tumour volume} = (\text{width})^2 \times \text{length}/2$ (ref. 25). The difference of tumour growth was statistically assessed by a MANOVA test²⁶.

Received 14 October; accepted 13 November 1996.

- Vitetta, E. S., Fulton, R. J., May, R. D., Till, M. & Uhr, J. W. *Science* **238**, 1098–1104 (1987).
- Pastan, I. & Fitzgerald, D. *Science* **254**, 1173–1177 (1991).
- Rosenberg, S. A. et al. *N. Engl. J. Med.* **319**, 1676–1680 (1988).
- Jain, R. K. *Sci. Am.* **271**, 58–65 (1994).
- Shockley, T. R. et al. *Cancer Res.* **52**, 357–366 (1992).
- Perez, P., Hoffman, R. W., Shaw, S., Bluestone, J. A. & Segal, D. M. *Nature* **316**, 354–356 (1985).
- Gross, G., Waks, T. & Eshhar, Z. *Proc. Natl Acad. Sci. USA* **86**, 10024–10028 (1989).
- Hwang, J., Fitzgerald, D. J., Adhya, S. & Pastan, I. *Cell* **48**, 129–136 (1987).
- Walter, P. & Lingappa, V. R. *Annu. Rev. Cell Biol.* **2**, 499–516 (1986).
- Kasprzyk, P. G., Song, S. U., DiFiore, P. P. & King, C. R. *Cancer Res.* **52**, 2771–2776 (1992).
- Batra, J. K., Kasprzyk, P. G., Bird, R. E., Pastan, I. & King, C. R. *Proc. Natl Acad. Sci. USA* **89**, 5867–5871 (1992).
- Chen, S.-Y., Khouri, Y., Bagley, J. & Marasco, W. A. *Proc. Natl Acad. Sci. USA* **91**, 5932–5936 (1994).
- Potter, C. R. D. et al. *Histopathol.* **15**, 351–362 (1989).
- Press, M. F., Cordon-Cardo, C. & Slamon, D. J. *Oncogene* **5**, 953–962 (1990).
- King, C. R., Fischer, P. H., Rando, R. F. & Pastan, I. *Sem. Cancer Biol.* **7**, 79–86 (1996).
- Slamon, D. J. et al. *Science* **244**, 707–712 (1989).
- Di Fiio, P. P. et al. *Science* **237**, 178–182 (1987).
- Kraus, M. H. et al. *EMBO J.* **6**, 605–610 (1987).
- Culver, K. et al. *Proc. Natl Acad. Sci. USA* **88**, 3155–3159 (1991).
- Miller, A. D., Trauber, D. R. & Buttimore, C. *Somat. Cell Mol. Genet.* **12**, 175–183 (1986).
- Yang, A. & Chen, S.-Y. *Nature Biotechnol.* (in the press).
- Marasco, W. A., Haseltine, W. A. & Chen, S.-Y. *Proc. Natl Acad. Sci. USA* **90**, 7889–7893 (1993).
- Collier, R. J. & Kandel, J. J. *Biol. Chem.* **246**, 1496–1503 (1971).
- Chen, S.-Y., Zani, C., Khouri, Y. & Marasco, W. A. *Gene Ther.* **2**, 116–123 (1995).
- Osborne, C. K., Hobbs, K. & Clark, G. M. *Cancer Res.* **45**, 584–590 (1985).
- Johnson, R. A. & Prentice-Hall, W. D. W. (eds) *Applied Multivariate Statistical Analysis* 226–290 (Englewood Cliffs, NJ, 1982).
- Miller, A. D. *Curr. Top. Microbiol. Immunol.* **158**, 1–24 (1992).

ACKNOWLEDGEMENTS. We thank Aronex Pharmaceuticals for purified recombinant immunotoxin and antibodies; H.-A. Fu for reagents; M. Miller, S. Akman, L. Kucera, A. Townsend and F. Torti for help and support; and S. Britt for secretarial assistance. This work was supported by institutional start-up funds, the Breast Cancer Pilot Project, an American Cancer Society institutional grant, and by the Cancer Research Foundation of America. S.-Y.C. is a recipient of an American Cancer Society junior faculty research award.

CORRESPONDENCE and requests for materials should be addressed to S.-Y.C. (e-mail: sichen@bgsu.edu).

P- and E-selectin mediate recruitment of T-helper-1 but not T-helper-2 cells into inflamed tissues

Frank Austrup*, Dietmar Vestweber†, Eric Borges‡, Max Löhning§, Rolf Bräuer§, Udo Herz||, Harald Renz||, Rupert Hallmann¶, Alexander Scheffold‡, Andreas Radbruch‡ & Alf Hamann*

* Abteilung für Immunologie, Medizinische Klinik, Universitätskrankenhaus Eppendorf, D-20246 Hamburg, Germany

† Institut für Zellbiologie, Universität Münster, D-48149 Münster, Germany

‡ Institut für Genetik, Universität Köln, D-50931 Köln, Germany

§ Institut für Pathologie, Klinikum des Friedrich-Schiller-Universität, D-07740 Jena, Germany

|| Institut für Klinische Chemie und Biochemie, Virchow-Klinikum des Humboldt Universität, D-13353 Berlin, Germany

¶ Institut für Experimentelle Medizin und Bindegewebforschung, Universität Erlangen-Nürnberg, D-91054 Erlangen, Germany

WHEN activated, T helper cells differentiate into one of two subsets, Th1 and Th2, characterized by distinct profiles of cytokine production. Th1 cells activate pro-inflammatory effector mechanisms involved in protection and autoimmunity, whereas Th2 cells induce humoral and allergic responses and downregulate local inflammation^{1,2}. Apart from differences in the repertoire of cytokines, no phenotypic attributes are established that distinguish the two subsets. Here we show that Th1 cells, but not Th2 cells, are able to bind to P-selectin and E-selectin. Moreover, only Th1 cells can efficiently enter inflamed sites in Th1-dominated models, such as sensitized skin or arthritic joints, but not in a Th2-dominated allergic response. Immigration of Th1 cells into inflamed skin can be blocked by antibodies against P- and E-selectin. These results provide evidence for adhesion mechanisms to distinguish between the two T helper subsets and mediate their differential trafficking. They indicate that selective recruitment is an additional level of regulation for both effector function profile and character of a local immune response.

The development of naive CD4⁺ cells into populations of Th1 and Th2 effector cells secreting unique patterns of cytokines can be directed by the presence of distinct cytokine combinations during activation^{3,4}. The acquired pattern of cytokine production has remarkable stability, even *in vivo*⁵, and determines the type of effector functions generated. Infection or other damage induces the local production of distinct cytokines by tissue cells and resident leukocytes; this probably initiates the differentiation of T cells reacting to the antigen *in situ* into either Th1 or Th2 cells. Within lymphoid tissues, however, the relative levels of cytokines might differ, resulting in the development of both kinds of Th cells. A second level of regulation should then maintain a polarized reaction within the local site. We therefore investigated whether Th1 and Th2 cells are differentially recruited into inflamed sites.

Purified CD4⁺ T cells were stimulated with anti-CD3 in the presence of either interleukin (IL)-4 and IL-2, or IL-12 and interferon (IFN)- γ . Alternatively, T cells transgenic for an ovalbumin-specific T-cell antigen receptor were stimulated by peptide⁶. As fully activated cells recirculate poorly⁷, the cells were allowed to return to a resting state before use. Most cells in the resulting populations produce cytokines when they are restimulated and these cells are exclusively of either Th2 or Th1 type, as shown by flow cytometric measurement of intracellular cytokine production of IL-4 and IFN- γ (Fig. 1a). The cells are small, essentially in phase G0 to G1, are low in IL-2 receptor, and have not yet acquired a typical memory phenotype, as expression of L-selectin and the CD45 splice variant RB is high (ref. 8 and

data not shown). The resulting populations probably resemble mature effector T cells *in vivo* in a stage at which they re-enter the circulation after experiencing a sessile, proliferative phase.

The migration of Th1 and Th2 cells into inflamed sites was compared in a cutaneous delayed-type hypersensitivity (DTH) reaction elicited by 2,4-dinitrofluorobenzene (DNFB). Striking differences in the immigration of Th1 and Th2 cells into the inflamed skin were found within 1 h of injection (Fig. 2a, b). Th1 cells entered these sites more than five times more efficiently than did Th2 cells; in a series of eight independent experiments, 1–2% of the transferred Th1 cells accumulated within 1 h in the 2.5-cm² DTH area compared with 0.1–0.3% of Th2 cells. The entry of unstimulated T cells or lymph-node-derived blasts was at the same low level (data not shown). Recovery of radioactivity was low in the non-inflamed control skin, and hardly differed between mice injected with Th1 or Th2, indicating that preferential immigration of Th1 cells into the cutaneous site is specific for the inflamed condition. Similar results were obtained with Th1 cells purified to >90% homogeneity using surface IFN- γ -recognizing liposomes⁶ (data not shown). Furthermore, injection of a 1:1 ratio of Th1 and Th2 cells, differentially labelled with intracellular fluorescent

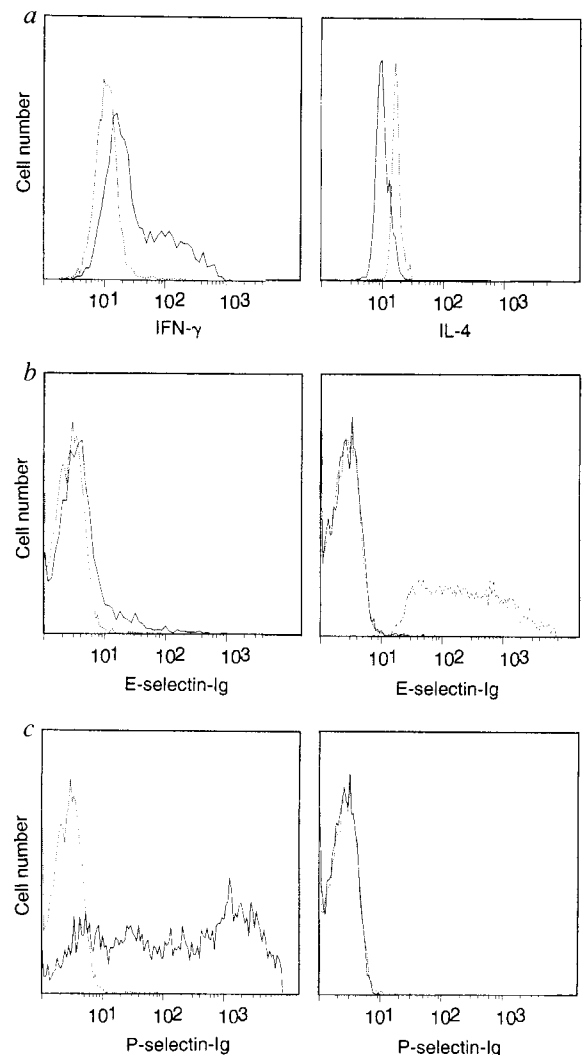


FIG. 1 Flow cytometry analysis of anti-CD3-stimulated Th1 and Th2 cells. a, Intracellular cytokine staining of Th1 (solid line) and Th2 cells (dotted line) for the cytokines IFN- γ and IL-4. b, c, Staining of Th1 (left) and Th2 cells (right) with E-selectin immunoglobulin (b) and P-selectin immunoglobulin (c) (solid lines) or a control immunoglobulin chimaeric protein (dotted line); a positive control for E-selectin binding (the myeloid cell 32Dcl3) is given (dot-dash line) in b, right.

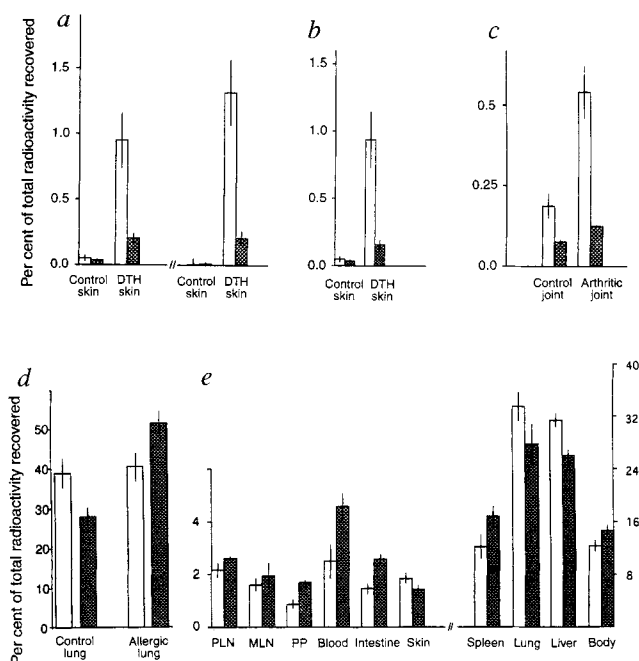


FIG. 2 *In vivo* migration of Th1 (open bars) and Th2 cells (hatched bars). *a*, Accumulation of radiolabelled anti-CD3-stimulated Th1 and Th2 cells in the inflamed skin 1 h (left) and 24 h (right) after intravenous injection. *b*, Immigration of ovalbumin peptide-stimulated Th1 and Th2 cells into inflamed skin (1 h after injection). *c*, Accumulation of anti-CD3-stimulated Th1 and Th2 cells in inflamed joints (3 h after injection). *d*, Localization of anti-CD3-stimulated Th1 and Th2 cells in the lung after induction of an allergic, asthma-like response (1 h after injection). *e*, Migration of anti-CD3-stimulated Th1 and Th2 cells into various tissues of untreated mice (1 h after injection). Note the different scale on the right part of the figure. Abbreviations: PLN, peripheral lymph nodes; MLN, mesenteric lymph nodes; PP, Peyer's patches.

dyes, resulted in a 7.5-fold higher accumulation of Th1 cells than Th2 cells in the sensitized ear after 1 h (data not shown). Th1 cells continued to accumulate slowly in the DTH region over a period of 24 h, whereas the level of Th2 cells remained low (Fig. 2a). Migration of Th1 cells was independent of how stimulation was achieved; in addition to the polyclonal activation by anti-CD3 (Fig. 2a) and the antigen-specific stimulation with peptide (Fig. 2b), allospecific Th1 and Th2 cells generated in a mixed lymphocyte reaction also showed the same difference (data not shown). A comparable bias for the immigration of Th1 cells into inflamed sites was found in cutaneous DTH reactions induced by the protein antigen keyhole limpet haemocyanin (data not shown).

Similar results were found in a murine, antigen-induced arthritis model. Th1 cells entered much more efficiently than Th2 cells

into the inflamed joints of animals primed with methylated bovine serum albumin and challenged intra-articularly 2 days before injection of the cells (Fig. 2c). A preferential immigration of Th1 cells into inflamed joints was also found at one, four and six days after challenge (data not shown). In contrast, eliciting of a Th2-dominated allergic response in a murine model of airway inflammation and hyper-responsiveness (U.H. *et al.*, submitted) did not increase recruitment of Th1 cells into the challenged lung (Fig. 2d).

No major differences in migration patterns were observed with respect to non-inflamed sites, except for the level of immigration into lung and liver, which was consistently higher with Th1 cells (Fig. 2e). Differences in blood levels and in the balance between lymphoid tissues versus lung and liver varied between experiments, and are likely to represent subtle differences in the differentiation/activation phase⁹.

These results indicate that Th1 cells are not only endowed with specific effector functions, but also have a distinct migration behaviour. The specific targeting of these cells into inflammatory sites of Th1-type reactions may thus increase their efficiency, as they are predominantly of use at the site of local immune reactions; this is in contrast to the systemic role of humoral responses stimulated by Th2-type help.

Immigration into human inflamed skin is thought to be mediated by sLe^x-related carbohydrate epitopes, called cutaneous lymphocyte antigen (CLA); CLA is recognized by E-selectin, which is strongly expressed in sites of cutaneous inflammation¹⁰. CLA was found to be induced *in vitro* by IL-12 or the transforming growth factor (TGF)- β ¹¹. We therefore investigated whether expression of E-selectin ligands is responsible for the preferential migration of Th1 cells into inflamed skin.

Staining with a soluble E-selectin-immunoglobulin construct¹², which heavily labels myeloid cells, revealed virtually no high-affinity E-selectin ligands on Th2 cells, and low levels on a minority of Th1 cells (Fig. 1b). In contrast, soluble P-selectin-immunoglobulin was found to stain most Th1 cells strongly, whereas Th2 cells remained unreactive (Fig. 1c). These expression patterns were found in both the polyclonally stimulated and the antigen-stimulated systems. P-selectin staining was Ca²⁺-dependent, and was blocked by the monoclonal antibody RB40.34 against P-selectin (data not shown).

To test whether expression of functional selectin ligands is essential for the selective recruitment into inflamed skin, we studied the immigration of Th1 cells under the influence of antibodies against P- and E-selectin. P-selectin antibodies only partly blocked the immigration of anti-CD3-stimulated Th1 cells into the DTH area (Fig. 3a). Despite the poor binding of soluble E-selectin to Th1 cells, antibodies against E-selectin also exerted a partial inhibitory effect. Moreover, a combination of antibodies against P- and E-selectin almost completely prevented Th1 cells entering the inflamed site. Thus P- and E-selectin are both essential for the recruitment of Th1 cells into a cutaneous inflammation. E- and P-selectin antibodies were not found to

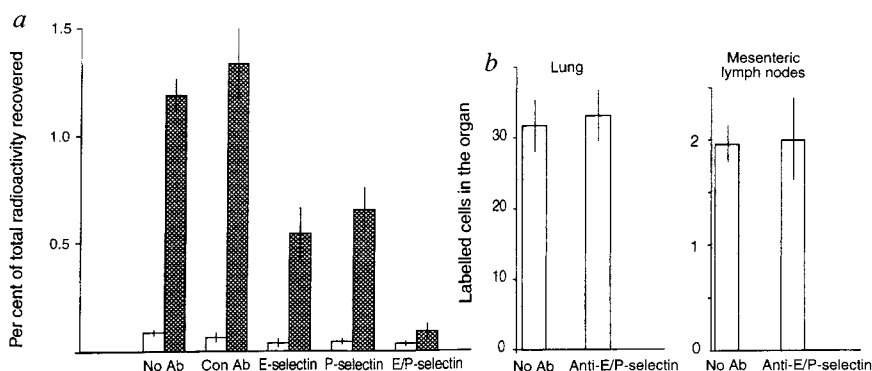


FIG. 3 Effects of antibodies (Ab) against the endothelial selectins. *a*, Inhibition of Th1 cell immigration into inflamed skin (hatched bars) or control skin (white bars) by antibodies against E- and P-selectin; *b*, A combination of E- and P-selectin antibodies had no effect on the localization of Th1 cells in lung and lymph nodes (and other organs, not shown). Radiolabelled anti-CD3-stimulated Th1 and Th2 cells were injected together with 200 μ g of each antibody. The percentage of labelled cells in the organs 1 h after injection is shown. Con Ab E, control antibody.

influence the localization of Th1 cells in non-inflamed organs, including the lung (Fig. 3b), which attracts considerable numbers of Th1 cells without challenge (Fig. 2d, e). In accordance with findings in other systems, which show a partly overlapping ligand specificity of P- and E-selectin, as well as overlapping function in neutrophil recruitment^{13,14}, our data demonstrate that both endothelial selectins aid the *in vivo* recruitment of T cells into the skin.

The apparent discrepancy between the contribution of E-selectin observed *in vivo* and the low reactivity of soluble E-selectin with Th1 cells *in vitro* suggests a possible role of low-affinity ligands for E-selectin, which are sufficient for transient bonds or acquire high avidity through multivalent binding *in vivo*. Indeed, a higher level of reactivity with E-selectin was found in flow cytometry experiments using either fluorescent latex beads coated with E-selectin (mean fluorescence: Th1 cells, 131; Th2 cells, 52; values for beads coated with P-selectin: Th1 cells, 668; Th2 cells, 88) or a liposome-based (multivalent) detection system (data not shown).

Thus the expression of functional selectin ligands distinguishes Th1 from Th2 cells, although we cannot exclude the possibility that a small percentage of Th1 cells might have no such ligands. The resulting difference in migration into sites of inflammation offers a level of regulation that targets distinct effector functions into tissue, and thereby contributes to the local balance between pro-inflammatory and suppressive activities. □

Methods

Generation and characterization of Th1 and Th2 cells. Resting, >98% pure CD4⁺ T cells were isolated by panning from pooled lymph-node lymphocytes of SPF-reared BALB/c mice⁸ using anti-CD8 (53-672), anti-CD25 (PC/6), anti-FcR II/III (2.4G2), anti-Mac-1 (M1/70) and anti-I-A^b (17/227) monoclonal antibodies. Cells (10⁶ per 1.5 ml) were activated on 24-well tissue-culture plates coated with an optimal amount (1 µg per well) of anti-CD3 monoclonal antibody (145-2C11) in the presence of the indicated recombinant murine cytokines (IL-2, 50 U ml⁻¹; IL-4, 10 ng ml⁻¹; IL-12, 1,000 U ml⁻¹; IFN-γ, 200 U ml⁻¹) for 2 days and then transferred without medium change onto blank plates for a further 4 days. In some experiments, 1 µg ml⁻¹ of anti-IL-4 or anti-IFN-γ antibody was added in the Th1 and Th2 cell cultures, respectively. Alternatively, Th1 and Th2 cells were generated from OVA-TCR transgenic mice¹⁵ as described⁶. For flow cytometric analysis of cytokine production, cells were washed, incubated for 24 h in medium without cytokines, and restimulated with TPA (50 ng ml⁻¹) and ionomycin (500 ng ml⁻¹) for 5 h in the presence of brefeldin A (10 µg ml⁻¹). The cells were fixed, permeabilized and stained¹⁶. As a negative control, cells were stained in the presence of an excess of soluble cytokines; the resulting background staining was almost identical to that of the non-producing cell type (Th1 cells or Th2 cells in the staining for IL-4 or IFN-γ, respectively). Cells without restimulation were negative, indicating that the cytokine added during culture was not carried over. The presence of ligands for E- or P-selectin was tested using soluble selectin-immunoglobulin chimerae (40 µg ml⁻¹), followed by PE-labelled F(ab)₂ anti-human immunoglobulin. A human CEA-immunoglobulin chimeraic protein was used as the negative control. The myeloid cell line 32Dcl3 strongly reacting with E- and P-selectin-immunoglobulin^{13,17} was used as a positive control.

Elicitation of inflammation and *in vivo* migration assays. A cutaneous DTH reaction was induced by skin painting with 0.5% 2,4-dinitrofluorobenzene (DNFB) in acetone-olive oil (4:1) on day -21, -20 and challenge with 0.3% DNFB at day -1. Antigen-induced arthritis was induced¹⁸ by subcutaneous priming of BALB/c mice with methylated BSA in complete Freund's adjuvant at day -21, a boost at day -14, and intra-articular challenge with methylated BSA at day -2. A Th2-dominated immune response in lung and airways with development of increased airway responsiveness was elicited by priming with ovalbumin/Al(OH)₃ intraperitoneally at day -22, -15, -8, followed by intranasal bacterial superantigen (SEB) stimulation at day -7, -5, -3 and bronchial challenge with aerosolized ovalbumin at day -2 and -1 as described (U.H., submitted). Th1 and Th2 cell populations were injected at day 0, that is, one day (DTH, lung hyper-responsiveness) or two days (arthritis) after challenge.

For analysis of lymphocyte migration^{19,20}, 2 × 10⁶ Th1 or Th2 cells per mouse were labelled with 20 µCi ⁵¹Cr per ml and injected into the tail vein of BALB/c mice (n = 4). Radioactivity in skin pieces of 2.5 cm² size, excised knee joints, various other organs and the remaining body was measured. Peripheral lymph nodes consisted of a pool of the superficial inguinal nodes, the brachial and axillary nodes, and the superficial cervical nodes, but did not include the mesenteric lymph nodes. All Peyer's patches were collected and counted separately from the remaining gut. Blood values were computed for a volume of 2 ml. Extended counting times and multiple background measurements were

used allowing a statistical error below 10% for samples containing down to 0.1% of radioactivity recovered within the animal (70–80% of injected radioactivity was recovered after 24 h). Similar results were obtained when Th1 or Th2 cells were prepared from either mesenteric or subcutaneous lymph nodes instead of using pooled lymph nodes (data not shown). As an alternative to radioactive measurements, T cells were labelled with the fluorescent dyes 5-chloromethyl-fluorescein diacetate or 5-(and-6)-(((4-chloromethyl)benzoyl)amino)tetramethyl-rhodamine (Molecular Probes) at a concentration of 4 µM (ref. 19).

For *in vivo* blockade of endothelial selectins, Th1 cells were co-injected with anti-E-selectin monoclonal antibody UZ4 (ref. 21) (200 µg per mouse), anti-P-selectin RB40.34/4 (ref. 22) (200 µg per mouse) or a combination of both. These antibodies are specific for the respective endothelial selectin, and do not crossreact with L-selectin. Indirect effects are excluded by the use of control antibodies, non-blocking anti-E-selectin antibodies, or further isotype controls (ref. 20 and data not shown).

Received 28 August; accepted 25 October 1996.

1. Mosmann, T. R. & Coffmann, R. L. *Annu. Rev. Immunol.* **7**, 145–173 (1989).
2. Mosmann, T. R. & Sad, S. *Immunol. Today* **17**, 138–146 (1996).
3. Fiorentino, D. F., Bond, M. W. & Mosmann, T. R. *J. Exp. Med.* **170**, 2081–2095 (1989).
4. Swain, S. L. *et al. Immunol. Rev.* **123**, 115–144 (1991).
5. Swain, S. L. *Immunity* **1**, 543–552 (1994).
6. Assenmacher, M. *et al. Eur. J. Immunol.* **26**, 263–267 (1996).
7. Smith, M. E., Martin, A. F. & Ford, W. L. *Monogr. Allergy* **16**, 203–232 (1980).
8. Austrup, F., Rebstock, S., Kilshaw, P. J. & Hamann, A. *Eur. J. Immunol.* **25**, 1487–1491 (1995).
9. Hamann, A. & Rebstock, S. in *Adhesion in Leucocyte Homing and Differentiation* (eds Dunon, D., C.R.M. & Imhof, B.) 109–124 (Springer, Berlin, 1993).
10. Berg, E. L. *et al. J. Exp. Med.* **174**, 1461–1466 (1991).
11. Leung, D. Y. *et al. J. Exp. Med.* **181**, 747–753 (1995).
12. Hahne, M., Jager, U., Isenmann, S., Hallmann, R. & Vestweber, D. *J. Cell Biol.* **121**, 655–664 (1993).
13. Lenter, M., Levinovitz, A., Isenmann, S. & Vestweber, D. *J. Cell Biol.* **125**, 471–481 (1994).
14. Labow, M. A. *et al. Immunity* **1**, 709–720 (1994).
15. Murphy, K., Heimberger, A. B. & Loh, D. Y. *Science* **250**, 1720–1723 (1990).
16. Assenmacher, M., Schmitz, J. A. & Radbruch, A. *Eur. J. Immunol.* **24**, 1097–1101 (1994).
17. Levinovitz, A., Muhlhoff, J., Isenmann, S. & Vestweber, D. *J. Cell Biol.* **121**, 449–459 (1993).
18. Bräuer, R., Thoss, K. & Henzgen, S. *Agents Actions* (suppl.) **39**, 69–73 (1993).
19. Hamann, A. & Jonas, P. in *Immunology Methods Manual* (ed. Lefkowitz, I.) 1333–1341 (Academic, London, 1996).
20. Hamann, A., Andrew, D. P., Jablonski-Westrich, D., Holzmann, B. & Butcher, E. C. *J. Immunol.* **152**, 3282–3293 (1994).
21. Hallmann, R., Zimmermann, U., Sorokin, L. M., Needham, L. & Mark, K.v.d. *Scand. J. Rheumatol.* **24** (suppl. 101), 107–109 (1995).
22. Bosse, R. & Vestweber, D. *Eur. J. Immunol.* **24**, 3019–3024 (1994).

ACKNOWLEDGEMENTS. We thank M. Gately and A. Stern for supplying IL-12; E. Schmitt and M. Assenmacher for help with the construction and characterization of the Th1/2 system; and D. Jablonski-Westrich and U. Zimmermann for technical assistance. This work was supported by the Mildred Scheel-Stiftung (A.H.), the Deutsche Forschungsgemeinschaft (A.H., D.V., U.H., H.R. and R.V.), and the Bundesministerium für Bildung, Wissenschaft, Forschung und Technologie (R.B.).

CORRESPONDENCE and requests for materials should be addressed to A.H. (e-mail: hamann@uke.uni-hamburg.de).

Selective modulation of protein kinase C-θ during T-cell activation

Colin R. F. Monks*†, Hannah Kupfer*, Idan Tamir*, Avlin Barlow‡ & Abraham Kupfer*†

* Division of Basic Sciences, National Jewish Center for Immunology, 1400 Jackson Street, Denver, Colorado 80206, USA

† Department of Cellular and Structural Biology, University of Colorado Health Sciences Center, Denver, Colorado 80262, USA

‡ Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510, USA

EVERY cell contains many families of protein kinases, and may express several structurally related yet genetically distinct kinases of each family. The activity of the serine/threonine protein kinase C (PKC) enzymes^{1–2} has long been implicated in T-cell activation³, but it is not known which members of the PKC family regulate the T-cell response to foreign antigens. The activation of T cells by antigen-presenting cells (APCs) is spatially restricted to their site of contact, where receptors on the T cells engage their counter-receptors on the APCs^{4,5}. We used this

localized engagement to identify, at the single-cell level, intracellular proteins involved in the activation process. By digital immunofluorescence microscopy, we localized six isoforms of PKC in antigen-specific T-cell clones activated by APCs. Surprisingly, only PKC- θ translocated to the site of cell contact. Accordingly, *in vitro* kinase activity assays of PKC immunoprecipitates from the conjugates of T cells and APCs showed a selective increase in the activity of PKC- θ , indicating that the translocated enzyme is active. Several modes of partial T-cell activation that failed to cause PKC- θ translocation also failed to cause T-cell proliferation, further suggesting the involvement of PKC- θ in T-cell activation.

Antigen-specific Th-APC conjugates were formed by mixing cloned Th D10 cells with CH12 cells that were pulsed with the antigen conalbumin. Figure 1 shows the immunofluorescence labelling for each of the following PKC isoforms: α , β_1 , δ , η , θ and ζ . All of these PKC isoforms are expressed in lymphocytes^{1,2} and were detected by digital microscopic analysis of Th-APC conjugates (Fig. 1). Surprisingly, only PKC- θ , (Figs 1a, b and 2a, b) was clustered at the cell membrane along the Th-APC contact site. The labelling for the other PKCs was distinctly punctate (Fig. 1c-l), and remained similar in both bound and unbound T cells (for example, PKC- δ in Fig. 2c-d). No clustering of any of the other PKC isoforms was detected in D10-CH12 conjugates after either shorter or longer times of interaction (data not shown).

PKC- θ shares a high sequence homology with PKC- δ (refs 6, 7), but its function is unknown. Unlike the other isoforms, PKC- θ has limited tissue distribution and has been reported to be predominantly expressed in haematopoietic cells⁷ and skeletal muscle⁶. Cell fractionation analysis of skeletal muscle⁸ indicates that, when treated with either TPA (12-*O*-tetradecanoylphorbol-13-acetate) or insulin, PKC- θ becomes associated with membrane, as do all of the other PKC isoforms tested. Similar results were reported with TPA activation of Jurkat T cells². Thus, with the exception of its selective expression, neither analysis of primary structure⁷ nor biochemical studies^{2,8} have previously indicated a unique role for PKC- θ in either skeletal muscle or T cells.

To demonstrate that the observed response was receptor mediated, and not a consequence of random cell-cell interactions, we tested the antigen specificity of the PKC- θ translocation. D10 cells were mixed with either untreated CH12 cells or with CH12 cells that were pulsed with conalbumin. The cells were then double labelled with anti-talin and anti-PKC- θ (Fig. 3), and analysed by digital microscopy. The clustering of talin, a cytoskeletal protein, at the Th-APC cell contact indicates, at the single-cell level, that the cells were activated^{4,9}. Within 10 min of cell mixing, most (70%) of the T cells that were bound to antigen-pulsed CH12 cells and clustered talin also displayed clustered PKC- θ at the cell contact (Fig. 3). This PKC- θ translocation was persistent and, by 55 min after cell mixing, PKC- θ was clustered in nearly all (88%) of the cell conjugates (Fig. 3). A three-dimensional quantitative analysis indicated that PKC- θ clustered in the D10 cells, but not in the CH12 cells (data not shown). In the absence of antigen, fewer cell conjugates were formed, and none of them clustered PKC- θ (Fig. 3). In similar studies, cells were also labelled for the other PKCs, and again there was no difference in their distribution in either antigen-specific or nonspecific conjugates (data not shown; as in Fig. 1). The clustering of PKC- θ was observed in each of eight other normal T-cell clones and with primary lymph-node T cells from T-cell antigen receptor (TCR)-transgenic mice, which were activated by either antigen or superantigen (data not shown). In contrast, the translocation of PKC- θ was not seen in any antigen-specific T-cell hybridomas (data not shown).

If PKC- θ is involved in T-cell activation, as would be predicted from its selective antigen-dependent translocation to the cell contact, it is likely to be enzymatically active in the T-APC conjugates. To measure its activity, PKC- θ was immunoprecipitated from mixtures of D10 and CH12 cells that were either untreated or pulsed with conalbumin. The immunoprecipitates

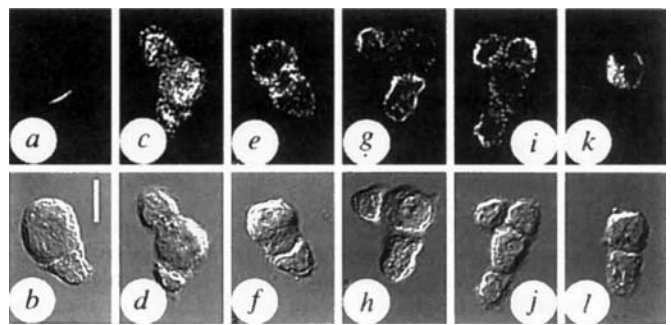


FIG. 1 Immunofluorescent localization of PKC isoforms in antigen-specific Th-APC conjugates. Top, the D10-CH12 conjugates were labelled with the following primary antibodies: rabbit anti-PKC- θ (a), PKC- ζ (c), PKC- η (e), PKC- δ (g), PKC- $\beta.1$ (i) and PKC- α (k); bottom, b-l, corresponding Nomarski images of the labelled cells shown above. The scale bar in b is 10 μ m.

were analysed in an *in vitro* kinase assay using the Selectide peptide¹⁰ as a substrate (Fig. 4). The *in vitro* kinase activities of PKC- θ from antigen-pulsed and unpulsed T-APC conjugates were strikingly different. PKC- θ from the cell mixtures that were not pulsed with antigen was largely inactive. The addition of TPA in the assay caused a major increase (430%) in the phosphorylation of the substrate. In contrast, PKC- θ from the antigen-pulsed cell mixtures was largely active, and the addition of TPA caused only a minor increase (13%) in the phosphorylation of the substrate. The similar activity of PKC- θ , in the presence of added TPA in the two immunoprecipitates indicates that the differences in the phosphorylation of the substrate reflect different states of activation, and not simply quantitative differences in the amount of immunoprecipitated PKC- θ (Fig. 4). Indeed, western blot analysis with anti-PKC- θ antibodies of the immunoprecipitates from the antigen-pulsed and unpulsed cell mixtures showed that roughly the same amounts of PKC- θ were present in the immunoprecipitates (Fig. 4). Microscopic observations indicated that a large fraction of cellular PKC- θ became associated with the antigen-specific T-APC cell contacts, so these findings strongly suggest that the translocated PKC- θ was enzymatically active.

Immunoprecipitates of PKC- δ , a PKC isoform that did not translocate to the cell contact (Figs 1 and 2), and is the most

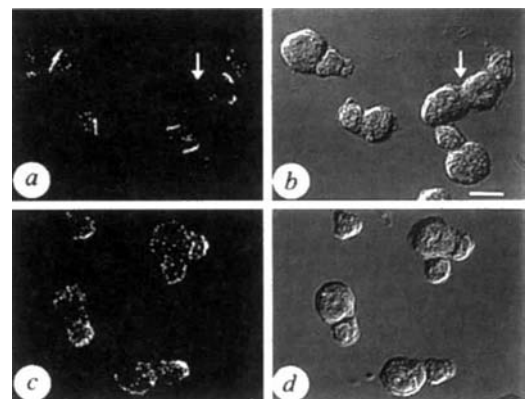


FIG. 2 The antigen-specific T-APC interactions selectively affect PKC- θ and not PKC- δ . Larger fields of view showing multiple D10-CH12 conjugates. The cells were labelled with either anti-PKC- θ (a) or PKC- δ (c). Note that PKC- θ is clustered in each of the five D10-APC contacts, but not clustered in the APC-APC contact (large arrow). PKC- δ is not clustered in any of the four D10-APC contacts. Note also that there is no difference between the distribution of PKC- δ in the APC-bound D10 cells and the unbound D10 cell. b, d, Nomarski images corresponding to a and c.

homologous to PKC- θ , were analysed alongside those of PKC- δ (Fig. 4). The *in vitro* kinase activity of PKC- δ , like its intracellular localization, was not affected by the interaction of the T cells with the antigen-pulsed APCs, and it remained largely inactive.

The selective TCR-dependent translocation of PKC- θ to the membrane site where the Th cell is bound to the APC suggests that PKC- θ is associated with other proteins that cluster at this site. Such target proteins may be engaged receptors, PKC substrates, or receptors for activated C kinase^{11,12}. Previous studies of antigen-specific Th-APCs showed that the TCR, CD4 and LFA-1 cluster, along with the cytoskeletal protein talin, at the cell contact^{9,13,14}. Antibody-induced capping experiments demonstrated that talin selectively associates with LFA-1 in activated T cells¹⁵. To identify receptors with which PKC- θ may be associated, TCR/CD3, CD4, LFA-1 and CD28 were capped with specific monoclonal antibodies in the absence of APCs. Digital microscopy revealed that PKC- θ did not colocalize with any of the capped receptors, including the capped TCR, regardless of the duration of the capping (data not shown). Additional co-capping experiments with anti-CTLA-4, CD43, CD45 and Thy-1 also failed to induce co-capping of PKC- θ (data not shown). Thus, although PKC- θ translocates, like talin, to the cell contacts during Th-APC interactions, simulated activations with either individual or combinations of monoclonal antibodies (data not shown) failed to trigger a similar co-capping. It is possible that PKC- θ failed to cap because it associates with an untested receptor. Alternatively, the response of PKC- θ may require spatially and temporally coordinated set of multiple signals that can be provided only by properly engaged APCs.

The antigen-induced selective activation and translocation of PKC- θ suggests that PKC- θ plays an important role in T-cell activation. We therefore investigated the responses of T cells activated under conditions that either cause or fail to cause the translocation of PKC- θ . The cloned D10 IL2 cells were mixed with CH12 cells that were pulsed with either high (500 $\mu\text{g ml}^{-1}$) or low

(0.5 $\mu\text{g ml}^{-1}$) concentrations of conalbumin. The cells were triply labelled with anti-PKC- θ , anti-talin and anti-tubulin to visualize the microtubule-organizing centre (MTOC). It was previously shown that activation of Th cells by APCs, which were pulsed with optimal concentrations of antigen, causes both clustering of talin and reorientation of MTOC, and results in T-cell proliferation⁹. In contrast, activation of Th cells by APCs pulsed with suboptimal concentrations of the same antigen is sufficient to induce clustering of talin, but does not result in reorientation of MTOC or T-cell proliferation⁹. Microscopic analysis of the D10-IL2-CH12 conjugates demonstrated that, in nearly all the cell conjugates formed in the presence of high concentrations of antigen, both talin and the MTOC were rearranged towards the contact area, and in these cells PKC- θ was clustered at the contact area (Fig. 5). In contrast, in the cell conjugates that were formed in the presence of a low concentration of antigen, talin did cluster at the contact area, but the MTOC remained randomly oriented, and PKC- θ was not clustered at the contact area (Fig. 5). In additional dose-response experiments the APCs were pulsed with intermediate concentrations of antigen; the numbers of cells with reoriented MTOCs and clustered PKC- θ decreased with a decrease in antigen. Most significantly, over the entire range of antigen concentrations, all T cells with clustered PKC- θ (>99%) had their MTOCs reoriented towards the bound APCs. These experiments indicate a tight coupling between the induction of both MTOC reorientation and cell proliferation, and PKC- θ translocation.

A similar correlation between the induction of PKC- θ translocation and cell proliferation was observed when we compared the responses of T cells to APCs pulsed with agonist and altered peptides. The cloned Th cell line AK8 (ref. 16), like the D10 cell line, is specific to the antigens conalbumin and IAK. The sequence of the conalbumin peptide that AK8 cells recognize as antigen is HRGAIEWEGIESG (CA132-144) (ref. 17). The addition of CA132-144 (1 $\mu\text{g ml}^{-1}$) to AK8 cells in the presence of irradiated APCs caused extensive T-cell proliferation, as measured by thymidine incorporation. In contrast, the peptide CA132-144/R133H (CA2), which includes a single amino-acid substitution in position 2 of CA132-144, inhibited the CA132-144-induced proliferation of AK8 cells in a dose-dependent manner. Microscopic observations showed that in nearly all of the AK8-CH12 conjugates that were pulsed with the CA132-144 peptide, both talin (95% $\pm$ 1.2% of the cells) and PKC- θ (85% $\pm$ 10%) were clustered at the cell contact area. In contrast, in cell conjugates that were pulsed with CA132-144 and CA2, conditions that result in extensive inhibition of cell proliferation (~80%), talin was enriched in nearly all of the conjugates (77 $\pm$ 4%), but PKC- θ clustered in only a few of the conjugates (18% $\pm$ 1.6%). The pulsing of APCs with only CA2, which does not induce detectable

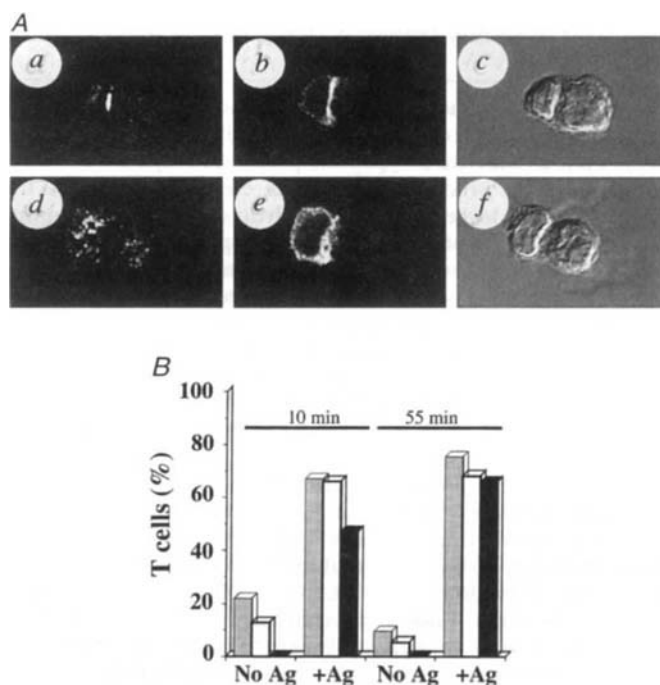


FIG. 3 The rapid and persistent translocation of PKC- θ is strictly antigen-specific. The D10-CH12 cell conjugates were formed with CH12 cells that were either incubated with conalbumin (a-c) or without conalbumin (d-f). A, Digital microscopy of D10-CH12 cell mixtures that were double labelled with affinity-purified rabbit anti-PKC- θ (a, d) and guinea-pig anti-talin (b, e) (ref. 5) antibodies. B, Effect of antigen (Ag) and time of interaction on the number of bound T cells, (grey), and T cells with clustered talin (white) or PKC- θ (black bars).

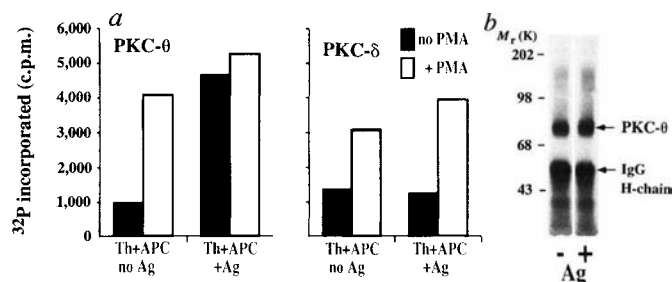


FIG. 4 *In vitro* kinase assays show that PKC- θ , but not PKC- δ , is activated during the interaction of T cells with antigen-pulsed APCs. a, Phosphorylation of the Selectide peptide by PKC- θ (left) and PKC- δ (right), which were immunoprecipitated from mixtures of T-APC that were incubated in the absence or presence of conalbumin. TPA, white bars; no TPA, black bars. b, Corresponding PKC- θ western blots of the PKC- θ immunoprecipitates. The PKC- θ activity is ~5-fold higher in the antigen-specific cell mixtures, but the amount of immunoprecipitated PKC- θ is almost identical.

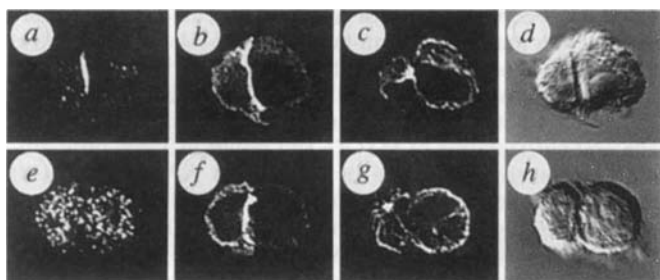


FIG. 5 Only APCs pulsed with high concentrations of antigen that induce full activation of T cells can cause the translocation of PKC- θ . CH12 cells were pulsed with 500 $\mu\text{g ml}^{-1}$ (a–d) or 0.5 $\mu\text{g ml}^{-1}$ (e–h) of conalbumin. D10–CH12 conjugates were triply labelled with anti-PKC- θ (a, e), talin (b, f) or tubulin (c, g). Note that PKC- θ is not clustered at the contact in the cell conjugate in which the MTOC is not facing the contact area.

proliferation of AK8 cells, resulted in the formation of significantly fewer Th–APC conjugates (about 20% of control levels). Talin was still clustered in most (70.5 $\pm$ 15%) of these conjugates, indicating some activation, but PKC- θ clustering was only very rarely (2% $\pm$ 1%) observed in these conjugates (data not shown).

The lack of PKC- θ translocation in antibody-activated T cells and in APC-activated T-cell hybridomas further supports the idea that PKC- θ participates in the induction of T-cell response to antigen. Treatment of D10 cells with anti-TCR/CD3 monoclonal antibodies induces cytokine production, but the cells later fail to proliferate and they undergo apoptosis (ref. 18, and C.R.F.M. *et al.*, unpublished data). Similarly, the activated T-cell hybridomas produce cytokines but fail to proliferate, and later undergo apoptosis^{19,20}. Taken together, these suggest that PKC- θ is involved in TCR-dependent proliferation, but not in triggering cytokine production. In summary, our studies at the single-cell level identified PKC- θ as being selectively affected during the interaction of T cells with their APC. Further studies are needed to identify the role and molecular basis for the unique response of this enzyme. □

Methods

Immunofluorescence microscopy. The cloned Th cell line D10.G4.1 (D10) is specific for the egg-white protein conalbumin and for IAK, and was used previously¹³. D10-IL2 is a subclone of D10. The B-cell lymphoma CH12.LX (CH12) (IAK) was used as an APC. D10 cells were mixed at a 1:1 cell ratio with CH12 cells that were pulsed overnight with conalbumin (500 $\mu\text{g ml}^{-1}$). Cell conjugates were formed and processed⁵. The affinity-purified rabbit antibodies against the PKC isoforms (Santa Cruz) were directed against unique peptide sequences at the C-terminal region of the proteins. Western blots with the anti-PKC- θ antibody labelled a single band (M_r ~80K) in GP&E fibroblasts that were infected with a retroviral vector encoding for PKC- θ , and did not detect any protein bands in the parental uninfected cells (data not shown). The ability of each of these antibodies to detect the specific PKC isoforms by immunofluorescence microscopy was confirmed by labelling, either in the presence or absence of TPA, a panel of fibroblast cell lines that overexpressed each of the PKC isoforms (data not shown). The immunofluorescence and the corresponding Normarski images of the cells were recorded by a chilled charge-coupled device digital camera (MCD1000, Spectra Source) that was mounted on a Zeiss Axiophot microscope, equipped with narrow-band optical filters (Chromatech). Images were processed by a no-neighbour deconvolution program to remove out-of-focus haze⁶. At least 400 cells were analysed per coverslip.

In vitro kinase assays. Conjugates of D10 (4×10^6 cells per assay) and CH12 (2×10^6 cells per assay) were formed as before, and 20 min later the cells were washed once in PBS and lysed (100 μl of 20 mM Tris, pH 7.6, and 0.5% NP-40, 0.25 M NaCl, 3 mM EDTA, 3 mM EGTA, 1 mM PMSF, 2 mM Na_2VO_4 , 20 $\mu\text{g ml}^{-1}$ Aprotinin, 100 $\mu\text{g ml}^{-1}$ Leupeptin, 1 mM DTT). After centrifugation, PKC- θ and PKC- δ were immunoprecipitated with the affinity-purified rabbit antibodies (1 μg) coupled to protein A Sepharose beads. The beads were washed in PKC-kinase buffer (20 mM HEPES, pH 7.2, and 137 mM NaCl, 5.4 mM KCl, 0.3 mM NaH_2PO_4 , 0.4 mM KH_2PO_4 , 25 mM β -glycerophosphate, 10 mM MgCl_2 , 5 mM EGTA and 2.5 mM CaCl_2). The kinase assay was performed by adding 35 μl of kinase buffer containing 0.1 mM ATP, [γ - ^{32}P]ATP (20 μCi per assay) and 0.1 mM Selectide PKC substrate (CalBiochem) for 20 min at 30 °C. The reactions were stopped by the addition of 10 μl 25% TCA. The reaction mixtures were blotted on

P81 phosphocellulose paper, washed with 75 mM phosphoric acid to remove free ATP, dehydrated and counted in a scintillation counter (Beckman). For each set of conditions, reactions were also performed in the absence of the Selectide substrate. The level of phosphorylation of the substrate was determined by subtracting the radioactive counts that were obtained in the absence of the substrate from those obtained in its presence. In each case a parallel reaction mixture contained also 125 ng ml^{-1} TPA, which causes maximal activation of the PKC that is present in the immunoprecipitate. Duplicate samples were analysed for each assay.

Received 9 August; accepted 8 November 1996.

- Hug, H. & Sarre, T. F. *Biochem. J.* **291**, 329–343 (1993).
- Baier, G. *et al.* *Eur. J. Biochem.* **225**, 195–203 (1994).
- Weiss, A. & Littman, D. R. *Cell* **76**, 263–274 (1994).
- Podack, E. R. & Kupfer, A. *Annu. Rev. Cell Biol.* **7**, 479–504 (1991).
- Kupfer, H., Monks, C. R. & Kupfer, A. *J. Exp. Med.* **179**, 1507–1515 (1994).
- Osada, S. *et al.* *Mol. Cell. Biol.* **12**, 3930–3938 (1992).
- Baier, G. *et al.* *J. Biol. Chem.* **268**, 4997–5004 (1993).
- Yamada, K. *et al.* *Biochem. J.* **308**, 177–180 (1995).
- Kupfer, A. & Singer, S. J. *J. Exp. Med.* **170**, 1697–1713 (1989).
- Chen, S. J. *et al.* *Biochemistry* **32**, 1032–1039 (1993).
- Robles-Flores, M. & Garcia-Sainz, J. A. *Biochem. J.* **296**, 467–472 (1993).
- Mochly-Rosen, D. *Science* **268**, 247–251 (1995).
- Kupfer, A. *et al.* *Proc. Natl Acad. Sci. USA* **84**, 5888–5792 (1987).
- Kupfer, A. & Singer, S. J. *Annu. Rev. Immunol.* **7**, 309–337 (1989).
- Kupfer, A., Burn, P. & Singer, S. J. *J. Mol. Cell. Immunol.* **4**, 317–325 (1990).
- Rojo, J. M., Kerner, J. D. & Janeway, C. Jr *Eur. J. Immunol.* **19**, 2061–2067 (1989).
- Yoon, S. T., Dianzani, U., Bottomly, K. & Janeway, C. A. Jr *Immunity* **1**, 563–569 (1994).
- Cherwinski, H. M., Semenuk, G. T. & Ransom, J. T. *J. Immunol.* **148**, 2996–3003 (1992).
- Harwell, L., Skidmore, B., Marrack, P. & Kappler, J. J. *J. Exp. Med.* **152**, 893–904 (1980).
- Yang, Y., Mercep, M., Ware, C. F. & Ashwell, J. D. *J. Exp. Med.* **181**, 1673–1682 (1995).

ACKNOWLEDGEMENTS. These studies were supported in part by grant from the NIH.

CORRESPONDENCE and requests for materials should be addressed to A.K. (e-mail: kupfer@njc.org).

RAIDD is a new ‘death’ adaptor molecule

Hangjun Duan & Vishva M. Dixit

Department of Pathology, University of Michigan Medical School, Ann Arbor, Michigan 48109, USA

THE effector arm of the cell-death pathway is composed of cysteine proteases belonging to the ICE/CED-3 family^{1,2}. In metazoan cells these exist as inactive polypeptide precursors (zymogens), each composed of a prodomain, which is cleaved to activate the protease, and a large and small catalytic subunit. The coupling of these ‘death’ proteases to signalling pathways is probably mediated by adaptor molecules that contain protein–protein interaction motifs such as the death domain¹. Here we describe such an adaptor molecule, RAIDD, which has an unusual bipartite architecture comprising a carboxy-terminal death domain that binds to the homologous domain in RIP, a serine/threonine kinase component of the death pathway^{3,4}. The amino-terminal domain is surprisingly homologous with the sequence of the prodomain of two ICE/CED-3 family members, human ICH-1 (ref. 5) and *Caenorhabditis elegans* CED-3 (ref. 6). This similar region mediates the binding of RAIDD to ICH-1 and CED-3, serving as a direct link to the death proteases, indicating that the prodomain may, through homophilic interactions, determine the specificity of binding of ICE/CED-3 zymogens to regulatory adaptor molecules. Finally, alternations in the sequence of the N-terminal domain that are equivalent to inactivating mutations in the *C. elegans ced-3* gene^{7,8} prevent homophilic binding, highlighting the potentially primordial nature of this interaction.

A sequence (IMAGE Consortium CloneID 109053) was identified in the NCBI GenBank expressed-sequence tag (EST) data base as having statistically significant homology ($P < 0.001$) to the prodomain of the human ICE-like protease ICH-1. This EST clone was used to isolate a full-length complementary DNA that encoded a protein of 200 amino acids with a predicted relative molecular mass of 22K (Fig. 1a). Given its sequence and interac-

Leu27 and Gly65 in the prodomain of CED-3, because these correspond to inactivating mutations of the *ced-3* gene in *C. elegans* designated 1,040 and 718 (refs 7, 8). Residues corresponding to Leu 27 and Gly 65 in CED-3 were also mutated in RAIDD and ICH-1: in all Leu-27 or Gly-65 mutants binding was prevented, highlighting the importance of these residues in the prodomain-NTD interaction (Fig. 3). Mutation of conserved hydrophobic residues in the ICH-1 prodomain, including residues Phe 82, Phe 85 and Leu 89 (mt7, mt5 and mt3, where mt means mutation), abolished binding to the NTD of RAIDD. Surprisingly, it took the mutation of both conserved negatively charged residues (Asp 83 and Glu 87; mt8) to eliminate binding (Fig. 3).

Having established the nature of the interactions of the NTD of RAIDD, we asked which death-domain-containing molecules were engaged by the death domain of RAIDD. We first examined all five known mammalian DD-containing proteins implicated in apoptosis⁹. As RAIDD did not bind *in vitro* to either of the two DD-containing receptors (Fas or TNFR-1) (data not shown), we investigated whether it could bind to the DD-containing-receptor-associated molecules FADD, TRADD or RIP. As shown in Fig. 4a, RAIDD specifically bound RIP but not FADD or TRADD *in vivo*. But in the presence of RIP, RAIDD was able to complex with TRADD (Fig. 4a). As RIP is a component of the TNFR-1 signalling complex⁴, it might recruit RAIDD to the complex; RAIDD, in turn, could recruit ICH-1 and so form a direct link to the effector ICE/CED-3-like death proteases. To test this idea, we transiently transfected 293 cells with expression vectors encoding TNFR-1, DD-containing adaptor molecules (TRADD, RIP, RAIDD, FADD), ICH-1, and ICE as a control (Fig. 4b, c). As anticipated, TNFR-1 complexed with RAIDD in the presence of TRADD and RIP (Fig. 4b), and through RAIDD, ICH-1 was recruited to the signalling complex (Fig. 4c). We next tested whether RAIDD, besides engaging the death pathway, could also activate NF- κ B, because its upstream partner RIP can participate in both pathways⁴. Transfected RAIDD did not activate an NF- κ B reporter construct (data not shown), consistent with a primary role in apoptosis. Our results indicate that RAIDD can function as an adaptor molecule in recruiting the death protease ICH-1 to the TNFR-1 signalling complex.

To address the function of RAIDD in apoptosis, MCF-7 cell lines were transiently transfected with an RAIDD expression

construct. These cells underwent apoptosis, which was inhibited by the ICE-peptide inhibitor z-VAD-fmk, a poxvirus-encoded serpin known to inhibit apoptosis, CrmA, and by catalytically inactive ICH-1, which was presumably behaving in a dominant-negative manner by competing out endogenous ICH-1 (Fig. 5). However, mutant ICH-1 or various putative dominant-negative versions of RAIDD were unable to block TNFR-1-induced cell death (data not shown), presumably because the alternative TRADD-FADD-FLICE death pathway was unimpeded.

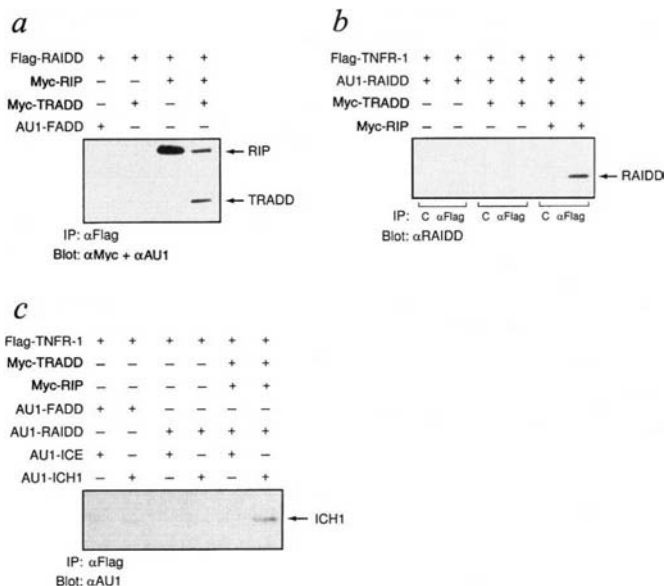


FIG. 4 On overexpression, RAIDD and ICH-1 are recruited to the TNFR1 signalling complex. a-c, 293 cells were transfected with the indicated combination of expression vectors. Detergent extracts of cells immunoprecipitated with either control mouse IgG or anti-Flag mAb. Co-precipitating proteins were analysed by immunoblotting with anti-epitope tag antibody (anti-Myc and anti-AU1 in a; anti-AU1 in c; or RAIDD antipeptide antibody in b).

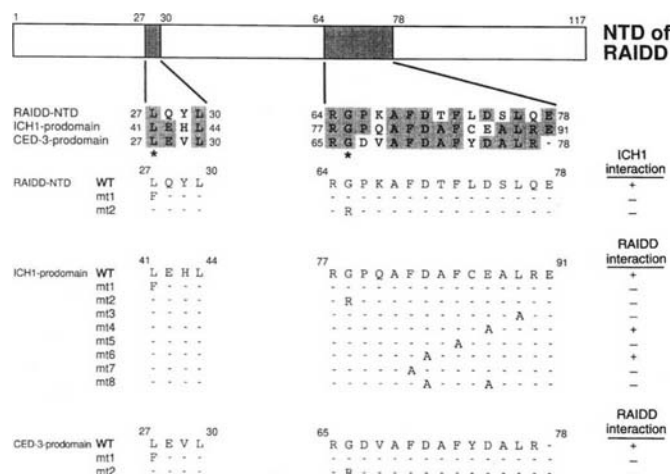


FIG. 3 Mutational analysis of the N-terminal domain in RAIDD and ICH-1. *In vitro* translated ³⁵S-labelled test proteins were incubated with either His-tagged RAIDD immobilized onto Ni²⁺ beads or ICH-1 proteins bound to protein-G-Sepharose. Alignment shows the regions of sequence similarity that surround the two known inactivating mutations in the prodomain of CED-3 (asterisk designated 1,040 and 718). Amino acids within each segment are numbered. Point mutations are shown as altered amino-acid residues aligned with the corresponding wild-type (WT) amino acid. A plus sign indicates significant binding; a minus sign indicates that binding was no different from background.

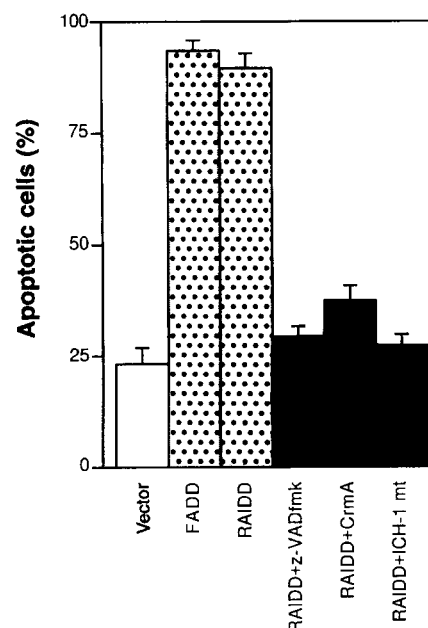


FIG. 5 RAIDD-induced apoptosis. Overexpression of RAIDD induces apoptosis that is inhibitable by CrmA, the broad-range ICE-family inhibitor z-VAD-fmk, and a catalytically inactive version of ICH-1 (Cys 302 $\rightarrow$ Ala). Data (mean $\pm$ s.e.m.) shown are the percentage of apoptotic cells among the total number of cells counted.

It is unclear why two seemingly redundant pathways should be deployed by a single receptor, but they may be used specifically by different cells or tissues. Our results have shown how prodomain interactions can impart binding specificity, a prerequisite for the assembly of a functioning death machine. □

Methods

cDNA cloning. Sequence corresponding to the prodomain of ICH-1 was used to screen the NCBI GenBank EST database using established methods. Several overlapping clones were identified as having statistically significant homology ($P < 0.001$) to the prodomain of ICH-1. One such clone, IMAGE Consortium CloneID 109053 (ref. 10), was characterized further and subjected to both automated and manual DNA sequencing. Additionally, a human K562 cell line cDNA library (from J. Lowe) was screened with a ^{32}P -labelled *PacI/EcoRI* restriction fragment of the EST clone¹¹. Hybridizing clones were characterized by automated sequencing. Sequence assembly, comparison and alignment were performed using DNASTAR software.

Northern blot analysis. Adult and fetal human multiple-tissue northern blots (Clontech) were hybridized according to the manufacturer's instructions with the same ^{32}P -labelled probe that was used for library screening.

Expression vectors. Mammalian cell expression vectors encoding AU1-FADD, Flag-TNFR1, Myc-TRADD and Myc-RIP have been described^{4,12}. AU1-RAIDD or Flag-RAIDD was generated by polymerase chain reaction (PCR) using custom-made oligonucleotide primers encoding the AU1 or Flag epitope. Amplified fragments were cloned into the mammalian expression vector pcDNA3 (Invitrogen). In-frame deletion mutants were generated by PCR with or without an N-terminal AU1 epitope and then subcloned into pcDNA3. Point mutants were created by site-directed mutagenesis using a two-step PCR method¹³ and confirmed by sequencing. A catalytically inactive ICH-1 was created by mutating the active-site Cys 302 to alanine.

In vitro binding assay. Constructs encoding His-tagged proteins were prepared in the prokaryotic expression vector pET23b (Novagen). Tagged proteins were expressed, purified and immobilized onto Ni^{2+} beads according to standard methodology. GST-Fas and -TNFR-1 fusion proteins were prepared as before¹². ^{35}S -labelled proteins were obtained by *in vitro* transcription/translation using a TNT T7-coupled reticulocyte lysate system (Promega). After translation, equivalent amounts of ^{35}S -labelled proteins were incubated with His-tagged proteins or GST fusion proteins immobilized onto Ni^{2+} beads or glutathione beads, respectively, as described^{12,14}. Beads were washed, boiled in SDS-sample buffer, and eluted proteins were resolved by SDS-PAGE and autoradiographed. Alternatively, ICH-1 protein was produced in human 293 embryonic kidney cells transfected with the AU1-ICH1 expression construct.

Transfection, coimmunoprecipitation and western analysis. 293 cells were transiently transfected with the indicated plasmids, lysed in 1 ml buffer (50 mM Tris, pH 7.6, 150 mM NaCl, 0.1% NP-40) and incubated either with mouse control IgG or anti-Flag mAb. Immune complexes were precipitated by the addition of protein-G-Sepharose (Sigma). After extensive washing, the Sepharose beads were boiled in sample buffer and the eluted proteins processed as before¹⁵.

Apoptosis assay. MCF-7 human breast carcinoma cells, CrmA-expressing stable cell lines¹⁶, or 293-EBNA cells were transiently transfected with 0.25 μg of the reporter plasmid pCMV- β -gal plus 1.5 μg of test plasmid encoding either FADD or RAIDD. The broad spectrum ICE-family inhibitor z-VAD-fmk (Enzyme Systems Products) was added to the cells at 20 μM , 5 h after transfection. 48 h later, apoptosis assay was as described¹².

Received 12 September; accepted 18 November 1996.

- Fraser, A. & Evan, G. *Cell* **85**, 781–784 (1996).
- Henkart, P. *Immunity* **4**, 195–201 (1996).
- Stanger, B. Z., Leder, P., Lee, T. H., Kim, E. & Seed, B. *Cell* **81**, 513–523 (1995).
- Hsu, H., Huang, J., Shu, H.-B., Baichwal, V. & Goeddel, D. V. *Immunity* **4**, 387–396 (1996).
- Wang, L., Miura, M., Bergeron, L., Zhu, H. & Yuan, J. *Cell* **78**, 739–750 (1994).
- Yuan, J., Shaham, S., Ledoux, S., Ellis, H. M. & Horvitz, H. R. *Cell* **75**, 641–652 (1993).
- Ellis, H. M. & Horvitz, H. R. *Cell* **44**, 817–829 (1986).
- Shaham, S. & Horvitz, H. R. *Genes Dev.* **10**, 578–591 (1996).
- Cleveland, J. L. & Ihle, J. N. *Cell* **81**, 479–482 (1995).
- Lennon, G., Auffray, C., Polymeropoulos, M. & Soares, M. B. *Genomics* **33**, 151–152 (1996).
- Sambrook, J., Fritsch, E. F. & Maniatis, T. *Molecular Cloning: A Laboratory Manual* 2nd edn (Cold Spring Harbor Laboratory Press, New York, 1989).
- Chinnaiyan, A. M., O'Rourke, K., Tewari, M. & Dixit, V. M. *Cell* **81**, 505–512 (1995).
- Higuchi, R., Krummel, B. & Saiki, R. K. *Nucleic Acids Res.* **16**, 7351–7367 (1988).
- Muzio, M. et al. *Cell* **85**, 817–827 (1996).
- O'Rourke, K. M., Laherty, C. D. & Dixit, V. M. *J. Biol. Chem.* **267**, 24921–24924 (1992).
- Tewari, M. & Dixit, V. M. *J. Biol. Chem.* **270**, 3255–3260 (1995).

ACKNOWLEDGEMENTS. We are especially grateful to H. R. Horvitz for ced-3 cDNA and D. V. Goeddel for RIP and TRADD expression constructs; we thank K. O'Rourke, A. M. Chinnaiyan and C. Vincenz for discussion and for material assistance, and I. Jones for help in preparing the figures.

CORRESPONDENCE and requests for materials should be addressed to V.M.D. (e-mail: vmcdixit@umich.edu). The nucleotide sequence reported here has been submitted to the GenBank/EMBL Data Bank (accession number U79115).

Brain acetylhydrolase that inactivates platelet-activating factor is a G-protein-like trimer

Yew S. Ho, Lora Swenson*, Urszula Derewenda, Laurence Serre*†, Yunyi Wei, Zbyszek Dauter‡, Mitsuharu Hattori§, Tomoya Adachi§, Junken Aoki§, Hiroyuki Arai§, Keizo Inoue§ & Zygmunt S. Derewenda

Department of Molecular Physiology and Biological Physics, University of Virginia Health Sciences Center, PO Box 10011, 1300 Jefferson Park Avenue, Charlottesville, Virginia 22906-0011, USA

* Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada

‡ EMBL Outstation, 22603 Hamburg, Germany

§ Department of Health Chemistry, Faculty of Pharmaceutical Sciences, The University of Tokyo, Hongo 7-3-1, Bunkyo-ku, Tokyo 113, Japan

THE platelet-activating factor PAF (1-*O*-alkyl-2-acetyl-*sn*-glycero-3-phosphocholine) is a potent lipid first messenger active in general cell activation, fertilization, inflammatory and allergic reactions, asthma, HIV pathogenesis, carcinogenesis, and apoptosis^{1–5}. There is substantial evidence that PAF is involved in intracellular signalling, but the pathways are poorly understood. Inactivation of PAF is carried out by specific intra- and extracellular acetylhydrolases⁶ (PAF-AHs), a subfamily of phospholipases A2 that remove the *sn*-2 acetyl group. Mammalian brain contains at least three intracellular isoforms, of which PAF-AH(Ib) is the best characterized^{7–9}. This isoform contains a heterodimer of two homologous catalytic subunits α_1 and α_2 , each of relative molecular mass 26K, and a non-catalytic 45K β -subunit, a homologue of the β -subunit of trimeric G proteins. We now report the crystal structure of the bovine α_1 subunit of PAF-AH(Ib) at 1.7 Å resolution in complex with a reaction product, acetate. The tertiary fold of this protein is closely reminiscent of that found in p21^{ras} and other GTPases. The active site is made up of a trypsin-like triad of Ser 47, His 195 and Asp 192. Thus, the intact PAF-AH(Ib) molecule is an unusual G-protein-like (α_1/α_2) β trimer.

Mammalian brain contains significant levels of PAF, which acts as a synapse messenger and transcription inducer of the early-response genes *c-fos* and *c-jun*¹⁰. It accumulates rapidly in neural tissue during seizures or ischaemia¹¹ and the resulting brain damage can be reduced by PAF antagonists¹². PAF is also being implicated as a messenger in long-term potentiation, a cellular model of memory formation¹³. The β -subunit of the PAF-AH(Ib) complex is a product of the gene that is responsible for the onset of type-1 lissencephaly, a developmental brain disorder caused by impaired neuronal migration, in which no cortex is formed¹⁴. This suggests a potential role for PAF in early brain development and neuronal migration.

The two catalytic α -subunits of PAF-AH(Ib) (this new notation is consistent with that accepted for the structurally related G-protein subunits) share extensive amino-acid sequence identity—63% in the bovine form⁹—but the consensus sequence they share is unique among proteins. When these subunits are purified or overexpressed individually in *Escherichia coli*, they form catalytically competent homodimers. *In vivo*, however, α_1 is preferentially labelled by [^{13}C]diisopropyl fluorophosphate in the heterodimer, suggesting functional asymmetry⁹.

The PAF-AH(Ib) α_1 subunit is a single polypeptide chain of 231 residues⁷. The molecule contains a single α/β domain with a central, parallel, 6-stranded β -sheet. This fold is very like that found in GTPases (Fig. 1a, 2). The insertions and deletions in the

† Present address: Institute de Biologie Structurale (IBS), Avenue des Martyrs, 38042 Grenoble Cedex, France.

α_1 structure, compared to p21^{ras}, can be rationalized in functional terms: for example, the turn/ β_1 fragment unique to p21^{ras} and involved in GTP binding is absent in α_1 , and the first β -strand found in the α -subunit of the G protein transducin and in p21^{ras}, which is involved in interactions with other proteins, is also unique to these proteins. On the other hand, α_1 contains a unique N-terminal α -helix which, together with the loop that follows, is part of the dimer-forming interface. Finally, the catalytic centre of α_1 is located on two Ω loops that are significantly different from their analogues in p21^{ras}.

Serine 47 has been identified as the putative nucleophile⁷. In the crystal structure, Ser 47 is close to His 195; further, the imidazole ring of His 195 is hydrogen-bonded through its N δ 1 atom to the side-chain carboxyl of Asp 192 (Fig. 1b). The stereochemistry of this triad is indicative of a catalytic function, which is supported by the inactivity of H195A and D192N mutants (data not shown). The chirality of the triad is the same as that found in the active sites of other esterases and neutral lipases, where nucleophilic attack is on the *re* face of the ester¹⁵ (the Cahn–Ingold–Prelog nomenclature specified the topologically different faces of peptides and esters involved in nucleophilic attack as *re* faces, not *re* and *si*¹⁵). The acetate bound in the active site interferes with the ground-state stereochemistry of the triad and disrupts the hydrogen bond that should link His 195 and Ser 47 in the active enzyme. The bound acetate also makes it possible to identify unambigu-

ously the hydrogen-donor groups that constitute the oxyanion hole: the main-chain amides of Ser 47 and Gly 74, and N δ 1 amide of Asn 104 (Fig. 1b). Cloned human plasma PAF-AH is also thought to work by means of a catalytic triad¹⁶.

The α_1 homodimer is formed by two molecules related by the crystallographic two-fold axis (Fig. 3a), oriented so that the carboxyl edges of their β -sheets are close to one another. In each monomer, a total of 1,150 Å² of solvent-accessible surface is lost when the dimer forms (Fig. 3b); 18 amino acids lose more than 20 Å² upon dimerization, and fourteen of these are conserved between the two catalytic subunits. Of the remaining four, only C55Y is a non-conservative substitution; this rationalizes the disorder between residues 53–56 (see Methods). Many of the interface-forming residues are hydrophilic and numerous hydrogen bonds and salt bridges mediate the contacts: for example, Gln 18–His 106, Asp 20–Arg 41, Ser 25–Gln 143 and Arg 29–Tyr 193; there are no buried water molecules within the interface.

The conserved character of the interface-forming residues indicates that the crystallographic dimer described here is a good model for the *in vivo* α_1/α_2 heterodimer. It is a doughnut-shaped structure with a solvent-accessible gorge, some 15 Å deep. The two active sites are at the bottom of the gorge and only 12 Å from each other, as measured between the carboxyl carbons of the acetate. The proximity suggests that these sites do not function independently, consistent with the noted functional asymmetry of

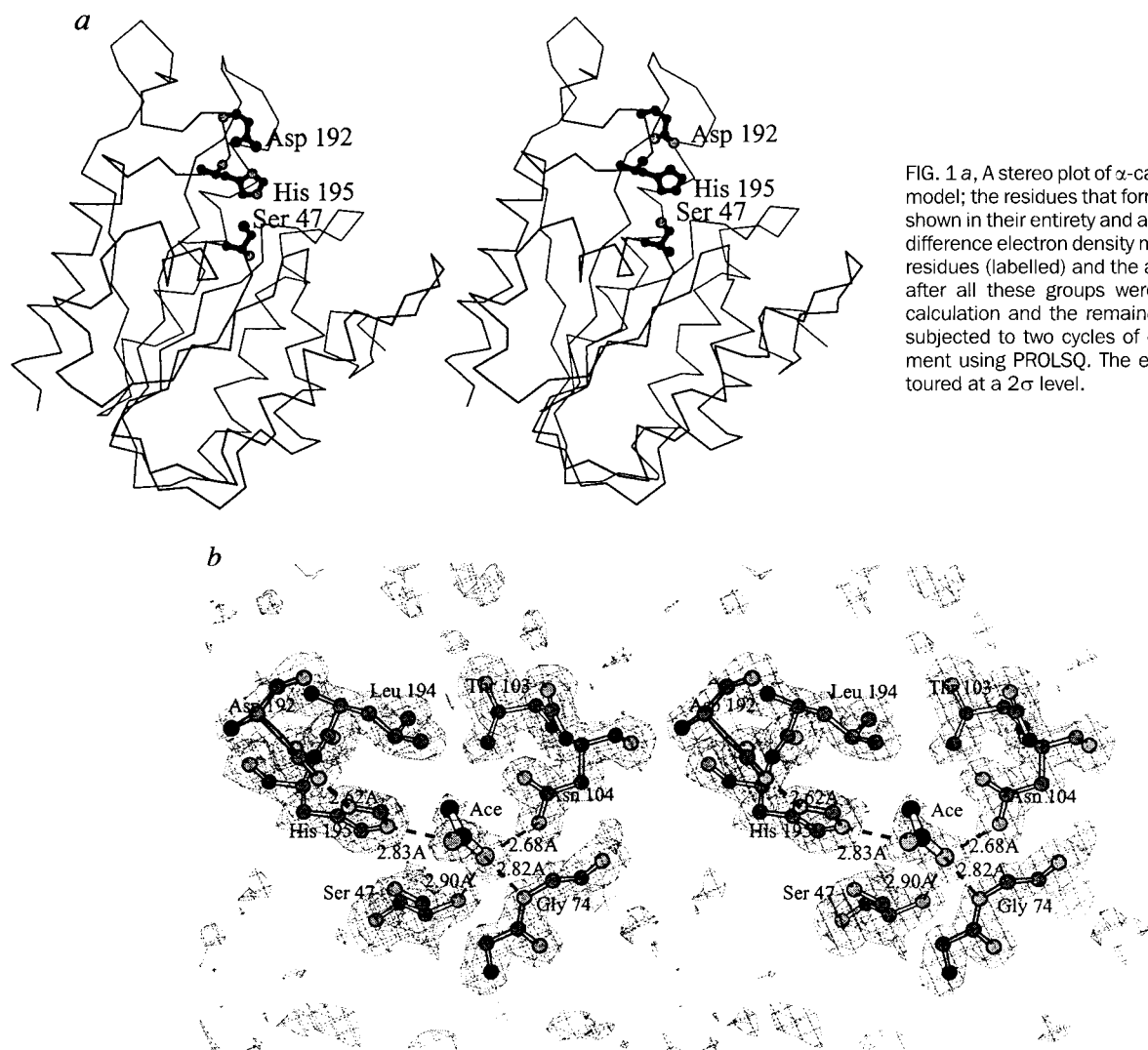


FIG. 1 a, A stereo plot of α -carbon atoms of the final model; the residues that form the catalytic triad are shown in their entirety and are labelled. b, An 'omit' difference electron density map showing active-site residues (labelled) and the acetate molecule (Ace) after all these groups were deleted from phase calculation and the remainder of the protein was subjected to two cycles of crystallographic refinement using PROLSQ. The electron density is contoured at a 2 σ level.

the heterodimer in which the activity of the α_2 subunit can be suppressed by steric effects.

In general catalysis by neutral lipases and secretory phospholipases A2 is markedly enhanced at substrate concentrations above their critical micellar concentrations (CMCs): plasma PAF-AH is significantly more active at PAF concentrations (CMC_{PAF}, 1.1 μ M) above 5 μ M (ref. 6). The structure of the α_1 homodimer shows that the entrance to the active gorge is not particularly hydrophobic, and there is no flexible secondary structure element that might function as a 'lid', a structural motif in serine-dependent lipases that are active at interfaces¹⁷. The crystal structure is therefore consistent with the enzyme being active on a monomeric substrate.

The activity of all PAF-AHs depends on their recognition of a

short acyl chain in the *sn*-2 position, but stringency varies between different isoforms. The plasma enzyme prefers the *sn*-2 acetyl group but will hydrolyse acyl chains of up to five carbons¹⁸; the brain isoform II also has broader specificity¹⁹. However, PAF-AH(Ib) is very selective, being 20 times less active towards propionyl-PAF than PAF itself¹⁹. The crystal structure can explain these effects. Three residues, all conserved in the α_2 subunit, come into contact with the acetate's methyl group: they are Leu 48, Leu 194 and Thr 103 (Fig. 4). Each donates one methyl group, that is, at C δ 1 (Leu 194), C δ 2 (Leu 48) and C γ 2 (Thr 103), to form a hydrophobic pocket for the substrate's acetyl moiety. Analysis of the probe-accessible surface created by these groups, and structural comparison with other neutral lipases indicate that the

FIG. 2 A comparison of the tertiary structures of PAF-AH(Ib) α_1 and GTPases: a, p21^{ras} (entry 5P21 in the Protein Data Bank); b, α_1 subunit of PAF-AH(Ib); c, catalytic domain of the transducin α -chain (entry 1GIA). Red, C α atoms used for least-squares overlap of PAF-AH(Ib) α_1 with either of the GTPases; green, the most significant insertions or deletions; blue, the partly disordered residues (55–61) in PAF-AH(Ib) α_1 . Figure generated using RIBBONS (M. Carson). Secondary structure elements of PAF-AH(Ib) α_1 : H0, residues 22 to 36; H1, 47 to 55; H2, 77 to 90; H3, 110 to 127; H4, 147 to 164; H5, 198 to 215; S1, 41 to 45; S2, 67 to 70; S3, 96 to 101; S4, 131 to 136; S5, 169 to 173.

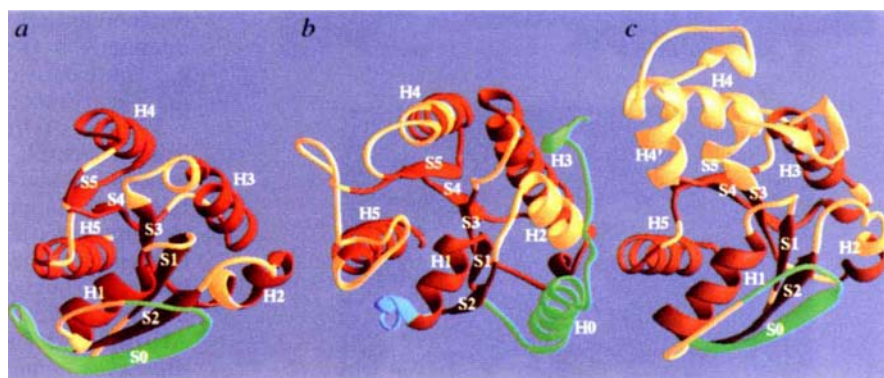


TABLE 1 Summary of crystallographic analysis

X-ray data collection	Native	Mercury	Lead	Gold	Cryo
Beam line (λ)	BW7B (0.895)	X31 (0.95)	X31 (0.93)	X31 (0.95)	BW7B (0.895)
Temperature	RT	RT	RT	RT	120 K
Total reflections	86,304	201,462	68,970	122,863	387,544
Unique reflections	17,228	15,029	13,983	15,006	33,716
Resolution range ($\text{\AA}$)	20–2.1	20–2.19	20–2.24	20–2.20	20–1.65
Completeness (%)	99.8	100	98.9	99.9	99.7
R_{sym} (%)*	8.4	8.2	9.1	7.3	6.3
R_{iso} (%)†	–	11.9	17.2	15.6	–
Overall phasing results					
R_{cullis} (acentric/centric)‡		0.83/0.76	0.81/0.76	0.88/0.85	
Phasing power (acentric/centric)§		1.20/0.90	1.30/0.80	0.80/0.50	
Number of sites (major/minor)		1/2	2/3	2/4	
Final refinement results (120 K) (data in parenthesis relate to a Δ 55–61 model)					
Resolution limit				8.5–1.7 $\text{\AA}$	
R_{factor} (%)				0.174 (0.195)	
Bond distance (1–2) r.m.s.				0.024 (0.015)	
Angle distance (1–3) r.m.s.				0.044 (0.034)	
Planar (1–4) r.m.s.				0.060 (0.047)	
Peptide planes r.m.s.				8.0° (5.7°)	
Main-chain average B -factor/r.m.s.				20.3 $\text{\AA}^2$ /1.4 $\text{\AA}^2$	
Side-chain average B -factor/r.m.s.				25.9 $\text{\AA}^2$ /3.1 $\text{\AA}^2$	
Residues in most favoured regions of Ramachandran plot¶				90.6%	
Residues in additional allowed regions				8.8%	

RT, room temperature

* $R_{\text{sym}} = \sum |I_i - \langle I \rangle| / \sum I_i$, where I_i is the intensity of the i th observation and $\langle I \rangle$ is the mean intensity of the reflection.

† $R_{\text{iso}} = \sum |F_{\text{PH}} - F_{\text{P}}| / \sum |F_{\text{P}}|$, where F_{P} and F_{PH} are the structure factor amplitudes for native and derivative crystals respectively.

‡ $R_{\text{cullis}} = \sum |\epsilon| / \sum |F_{\text{PH}} - F_{\text{P}}|$.

§ Phasing power = $\langle F_{\text{H}} \rangle / \langle \epsilon \rangle$, where $\langle F_{\text{H}} \rangle$ is the mean heavy-atom contribution and $\langle \epsilon \rangle$ is the mean lack of closure.

|| $R_{\text{factor}} = \sum |F_{\text{obs}} - F_{\text{calc}}| / \sum |F_{\text{obs}}|$.

¶ The analysis of the Ramachandran plot was done using PROCHECK²⁹.

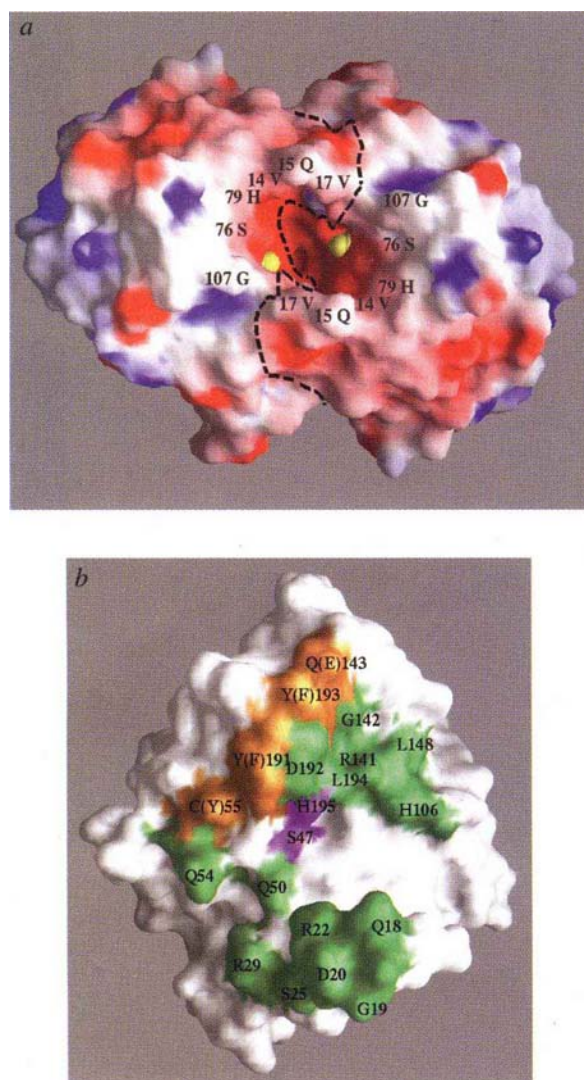
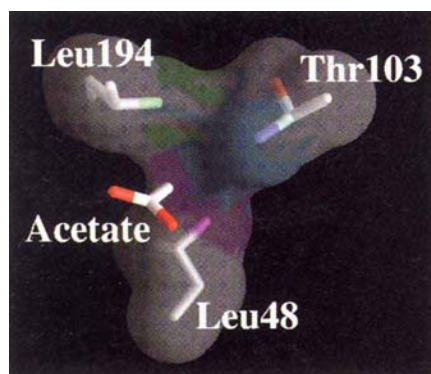


FIG. 3 *a*, Structure of the α_1 homodimer, looking down the entrance to the active-site gorge; black dotted line denotes the boundary between the monomers, and selected residues that line the entrance to the gorge are labelled. Bound acetate molecules are shown in yellow, and the surfaces accessible to probe are coloured according to their electrostatic potential. *b*, Dimer-forming interface of the α_1 subunit. The view here is of the right monomer shown in *a*, rotated 90° around the vertical axis. The residues indicated are those that lose at least $20 \text{ \AA}$ of solvent-accessible surface upon dimerization, as well as catalytic Ser 46 and His 195 (magenta), shown for clarity. They are coloured as follows: amino acids conserved between α_1 and α_2 are in green; non-conserved residues are shown in gold. Figure generated by GRASP³⁰.



methyl of Thr 103 is critical.

The different specificities of PAF-AHs may be important. Membrane phospholipids can be oxidized to molecules containing shorter, frequently oxidized groups in the *sn*-2 position, and these molecules can mimic the cell-damaging effects of PAF. Hence, PAF-AHs with broader specificity and preference towards insoluble substrates—in particular the plasma isoform—may act as general scavengers of PAF and PAF-like inflammatory phospholipids. Conversely, enzymes with high specificity, like PAF-AH(Ib), respond selectively to PAF, as would be expected if they form part of a signal transduction pathway.

The β -subunit of PAF-AH(Ib) shows limited but significant sequence homology with the β -subunits of G proteins, one of which—transducin G_{ip} —has the tertiary fold of a seven-blade β -propeller²⁰. Sequence alignment of the two proteins (not shown) indicates that, except for the N-terminal fragments—nearly 100 residues in the β -subunit and 46 in G_{ip} —both proteins share a common tertiary scaffold. Thus, the intact PAF-AH(Ib) trimer is essentially a G-protein-like (α_1/α_2) β molecule; the β -subunit appears to be associated exclusively with the α_1 chain²¹, supporting the idea of a functional asymmetry for the catalytic heterodimer.

The PAF-AH(Ib) β -subunit is important in brain development, being selectively expressed in the developing brain in regions undergoing neuronal migration in the cortex and cerebellum^{22,23}; this pattern supports a role for this subunit in neuronal migration. Its importance is underscored by its remarkable conservation: there is only one substitution among its 410 amino acids in the bovine and human species⁸, and only one between the murine and human forms²². There is one known fungal homologue of PAF-AH(Ib) β —45% amino-acid identity—the *NudF* gene product of *Aspergillus nidulans*²⁴, which is essential for the movement of nuclei along the hyphae of filamentous fungi, a process requiring microtubules and cytoplasmic dynein. An evolutionarily ancient phenomenon, nuclear movement, may have been recruited in the developing brain to initiate neuronal migration through as-yet unidentified interactions of PAF-AH(Ib) β with the cytoskeleton, as suggested by the binding of WD-repeat proteins, including PAF-AH(Ib) β , to pleckstrin-homology (PH) domains, including those of β -spectrin and dynamin²⁵.

Why does the catalytic α_1/α_2 dimer associate with the β -subunit to form a heterotrimer? Messenger RNA expression of PAF-AH(Ib) α_1 indicates that in the human fetus the transcript is found predominantly in the brain, although it can be detected in several tissues in the adult²⁶. Hydrolysis of PAF may induce conformational changes in the trimer that affect the ability of the β -subunit to interact with cytoskeletal proteins. We are investigating the nature of the interactions between the subunits of the PAF-AH(Ib) complex. □

Methods

Purification of the enzyme has been described²¹. The protein was crystallized using $3 \mu\text{l}$ of the protein at $4.3\text{--}6.3 \text{ mg ml}^{-1}$ in 10 mM Tris-HCl, pH 7.5, 2 mM DTT, mixed 1:1 with well solution (16% saturated ammonium sulphate in 200 mM sodium acetate buffer, pH 6.5) and equilibrated in a hanging drop at room temperature. Trigonal crystals, space group $P3_121$, $a = b = 82.3 \text{ \AA}$, $c = 72.9 \text{ \AA}$, grew to $0.4 \times 0.5 \times 0.6 \text{ mm}$ after 3–4 weeks, giving a diffraction pattern extending to $\sim 2.5 \text{ \AA}$ on a rotating anode X-ray source. Hg-, Pb- and Au-derivatized crystals were prepared by transferring native crystals to solutions containing 1 mM thimerosal, 500 mM trimethyl lead acetate and 5 mM sodium gold acetate, respectively, and equilibrating for 36 h. Data were collected at the EMBL Outstation in Hamburg. Native data at room temperature were collected on the Wiggler beamline BW7B; all data for derivatized crystals were collected on

FIG. 4 The specificity pocket, as illustrated by probe($1.4 \text{ \AA}$)-accessible surfaces of amino acids Leu 48, Leu 194 and Thr 103. The three critical methyl groups are shown in purple, green and blue, respectively. Figure generated by GRASP³⁰.

the tunable bending magnet beamline X31 where the wavelength was adjusted to maximize the anomalous signal from the heavy atoms. In all experiments, only one crystal was needed for a complete set. A single native crystal was flash-frozen after being transferred to a cryoprotectant solution of 30% glycerol in the crystallization medium. The unit cell dimensions after freezing were $a = b = 81.3 \text{ \AA}$, $c = 72.6 \text{ \AA}$. Data were reduced and scaled using the HKL package containing DENZO and SCALEPACK programs (Z. Otwinowski and W. Minor). Unless specified otherwise, all software was used as implemented in the CCP4 suite of programs (Daresbury Laboratory, SERC, UK). Phases for room-temperature data were refined using MLPHARE and further enhanced using density modification, as implemented in SQUASH²⁷. The resulting map was of sufficient quality to allow 50% of the main-chain to be traced as polyaniline. This partial model was used in combination in SIGMA (R. Read) to improve the phasing. The structure was traced using program O (ref. 28) and the model was refined at 2.1 Å resolution using PROLSQ, initially with residues 5–219. Water molecules were then added and refinement was continued with 13 residues (53–65) deleted from the model owing to intrinsic disorder in that region. When the 1.7 Å data were introduced, the orientation of the molecule was optimized using AMORE (J. Navaza) and X-PLOR (A. Brünger) version 3.1; only a small deviation in rotation and translation parameters was observed; the resulting R -factor for this model against all data was 0.33. A parallel calculation was done using only 90% of data to monitor the R_{free} value for 10% of randomly chosen reflections. The 2.1 Å model was subjected directly to two rounds of slow cooling using X-PLOR, followed by PROLSQ and REFMAC refinement of positional and thermal parameters; no manual revision of the model was done. As a result, the R factor was reduced for a model including 141 water molecules to 0.22, while the R_{free} was 0.27. At this point, refinement using all data was resumed at the point at which it was left. Several cycles of PROLSQ reduced the R factor to below 0.20. A number of side chains were corrected manually in program O and the acetate, now clearly defined in the active site, was included in the refinement. Of the 13 residues deleted from the previous model, only 4 had to be modelled into ambiguous density. The C-terminal amino acids 216–219 were deleted owing to the absence of interpretable density. Solvent structure and the model (consisting of residues 5–216, an acetate ion bound in the active site, and 241 water molecules) were refined using REFMAC, a maximum-likelihood refinement program. Because of the partial disorder of residues 55–61, their stereochemistry was relatively poor. A final refinement calculation was therefore done for a $\Delta 55$ –61 model from which they were deleted. Stereochemical parameters for both models are listed in Table 1.

Received 3 September; accepted 4 November 1996.

- Venable, M. E., Zimmerman, G. A., McIntyre, T. M. & Prescott, S. M. *J. Lipid Res.* **34**, 691–702 (1993).
- O'Flaherty, J. T. & Wykle, R. L. in *Platelet Activating Factor and Human Disease* (eds Barnes, P. J., Page, C. P. & Henson, P. J.) 117–137 (Blackwell Scientific, Oxford, 1989).
- El Azzouzi, B., Jurgens, P., Benevise, P. & Thomas, Y. *Biochem. Biophys. Res. Commun.* **190**, 320–324 (1993).
- Shukla, D. *FASEB J.* **6**, 2296–2301 (1992).
- Bennett, S. A., Leite, L. C. & Birnboim, H. C. *Carcinogenesis* **14**, 1289–1296 (1993).
- Stafforini, D. M., Prescott, S. M., Zimmerman, G. A. & McIntyre, T. M. *Biochim. Biophys. Acta* **1301**, 161–173 (1996).
- Hattori, M., Adachi, H., Tsujimoto, M., Arai, H. & Inoue, K. *J. Biol. Chem.* **269**, 23150–23155 (1994).
- Hattori, M., Adachi, H., Tsujimoto, M., Arai, H. & Inoue, K. *Nature* **370**, 216–218 (1994).
- Hattori, M. et al. *J. Biol. Chem.* **270**, 31345–31352 (1995).
- Bazan, N. G. & Rodriguez de Turco, E. B. *Neurochem. Intern.* **26**, 435–441 (1995).
- Kumar, R., Harvey, S., Kester, N., Hanahan, D. & Olson, M. *Biochim. Biophys. Acta* **963**, 375–383 (1988).
- Panetta, T., Marcheselli, V. L., Braquet, P., Spinnewyn, B. & Bazan, N. G. *Biochem. Biophys. Res. Commun.* **149**, 580–587 (1987).
- Kato, K., Clark, G. D., Bazan, N. G. & Zorumski, D. F. *Nature* **367**, 175–179 (1994).
- Reiner, O. et al. *Nature* **364**, 717–721 (1993).
- Derewenda, Z. S. & Wei, Y. J. *Am. Chem. Soc.* **117**, 2104–2105 (1995).
- Tjoelker, L. W. et al. *J. Biol. Chem.* **270**, 25481–25487 (1995).
- Derewenda, Z. S. *Adv. Prot. Chem.* **45**, 1–52 (1994).
- Stremmel, K. E., Stafforini, D. M., Prescott, S. M. & McIntyre, T. M. *J. Biol. Chem.* **266**, 11095–11103 (1991).
- Hattori, K. et al. *J. Biol. Chem.* **270**, 22308–22313 (1995).
- Sondek, J., Böhm, A., Lambright, D. G., Hamm, H. E. & Sigler, P. B. *Nature* **379**, 369–374 (1996).
- Hattori, M., Arai, H. & Inoue, K. *J. Biol. Chem.* **268**, 18748–18753 (1993).
- Reiner, O. et al. *J. Neurosci.* **15**, 3730–3738 (1995).
- Mizuguchi, M., Takashima, S., Kakita, A., Yamada, M. & Ikeda, K. *Am. J. Path.* **147**, 1142–1151 (1995).
- Xiang, X., Osmani, A. H., Osmani, S. A., Xin, M. & Morris, N. R. *Mol. Biol. Cell* **6**, 297–310 (1995).
- Wang, D. S. et al. *Biochem. Biophys. Res. Commun.* **209**, 622–629 (1995).
- Adachi, H., Tsujimoto, M., Hattori, M., Arai, H. & Inoue, K. *Biochem. Biophys. Res. Commun.* **214**, 180–187 (1995).
- Zhang, K. Y. J. *Acta Crystallogr. D* **49**, 213–222 (1993).
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, M. *Acta Crystallogr. A* **47**, 110–119 (1991).
- Laskowski, R. A., McArthur, M. W., Moss, D. S. & Thornton, J. M. *J. Appl. Crystal.* **20**, 282–291 (1993).
- Nichols, A., Sharp, K. & Honig, B. *Proteins* **11**, 281–296 (1991).

CORRESPONDENCE and requests for materials should be addressed to Z.S.D. (e-mail: zsd4n@virginia.edu). Atomic coordinates have been deposited in the Brookhaven Protein Data Bank, entry code 1WAB, where they will be held for a year before release.

Regulatory intramolecular association in a tyrosine kinase of the Tec family

Amy Hamilton Andreotti*, Stephen C. Bunnell†, Sibong Feng*, Leslie J. Berg† & Stuart L. Schreiber*

* Howard Hughes Medical Institute, Department of Chemistry and Chemical Biology, Harvard University, 12 Oxford Street, Cambridge, Massachusetts 02138, USA

† Department of Molecular and Cellular Biology, Harvard University, 16 Divinity Avenue, Cambridge, Massachusetts 02138, USA

THE T-cell-specific tyrosine kinase Itk is a member of the Tec family of non-receptor tyrosine kinases^{1–7}, and is required for signalling through the T-cell antigen receptor (TCR)⁸. The role of Itk in TCR signalling and the manner in which Itk activity is regulated are not well understood. Substrate binding and enzymatic activity of the structurally related Src kinases are regulated by an intramolecular interaction between the Src-homology-2 (SH2) domain and a phosphotyrosine^{9,10}. Although Itk also contains SH3, SH2 and tyrosine kinase domains, it lacks the corresponding regulatory phosphorylation site, and therefore must be regulated by an alternative mechanism. The proline-rich sequence adjacent to the SH3 domain of Tec family kinases contains an SH3 ligand, potentially allowing a different intramolecular interaction. By using multidimensional nuclear magnetic resonance we have determined the structure of a fragment of Itk, confirming that these domains interact intramolecularly. Formation of this intramolecular SH3–ligand complex prevents the Itk SH3 domain and proline-rich region from interacting with their respective protein ligands, Sam68 and Grb-2. We believe that this structure represents the first example of an intramolecular interaction between an SH3 domain and a proline-rich ligand, and has implications for the regulation of Tec family kinases.

The amino-acid sequence of the proline-rich region, KPLPPTP, precedes the SH3 domain in Itk and conforms to the consensus of an SH3 ligand, (R/K)XXPXXP^{11–15}. Although most SH3 domains bind preferentially to Arg-containing peptides, screening of phage display peptide libraries with the Itk SH3 domain revealed a preference for Lys as the critical basic residue¹⁶. The sequence similarity between the phage display peptide and the native Itk sequence prompted us to investigate the possibility that the SH3 domain of Itk could interact with the Itk proline-rich region. However, it has not been possible to demonstrate an intermolecular interaction between the proline-rich peptides and SH3 domains of Tec family kinases (ref. 17, and A.H.A. and S.L.S., unpublished results). Because the location of a putative ligand for the Itk SH3 domain adjacent to the Itk SH3 domain suggested intriguing regulatory possibilities, we chose to determine whether these domains interact intramolecularly.

To address this issue, the Itk SH3 domain (referred to as ItkSH3) and a fragment containing both the SH3 domain and the potential ligand site (ItkSH3-KPL) were expressed in *Escherichia coli* and purified (Fig. 1a). Consistent with an intramolecular (as opposed to intermolecular) interaction, the oligomerization state of ItkSH3-KPL was found to be monomeric based on several methods using the monomeric ItkSH3 as a control. A three-dimensional half-filtered nuclear Overhauser enhancement spectroscopy (NOESY) experiment¹⁸, in which NOEs between ¹³C-attached protons and ¹²C-attached protons are correlated, was performed on a sample containing a 1:1 mixture of ¹³C-labelled and unlabelled ItkSH3-KPL, and revealed no intermolecular NOEs. A ¹³C-dispersed NOESY experiment^{19,20}, in which NOEs between ¹³C-attached protons are observed, does exhibit NOEs between ¹³C-attached protons of

the KLPPTP sequence and of the SH3 binding surface (A.H.A. and S.L.S., unpublished results). These data therefore suggest that the KLPPTP sequence and the SH3 domain within one ItkSH3-KPL molecule do not associate intermolecularly, and are consistent with an intramolecular model for SH3–ligand self-association. In further support of this model, elution times from an analytical size-exclusion column and nuclear magnetic resonance (NMR) linewidths, which are both indirect measures of apparent molecular weight, were similar for ItkSH3-KPL and ItkSH3. In addition, fluorescence perturbations caused by ligand-induced desolvation of the exposed tryptophan in the SH3-binding site have previously been shown to be concentration dependent for a fragment of the phosphatidylinositol 3-kinase p85 subunit that dimerizes in solution²¹. In contrast, the fluorescence intensity of ItkSH3-KPL increases in a concentration-independent manner consistent with an intramolecular association.

Determining the solution structure of ItkSH3-KPL by multi-dimensional NMR revealed a specific association between the SH3 domain and the proline-rich region of Itk. The structure of the intramolecular complex formed between the Itk SH3 domain and the adjoining ligand, KLPPTP, is shown in Fig. 1*b–d*. The SH3 domain of Itk adopts a fold that closely resembles other known SH3 structures^{13–15}. Two three-stranded anti-parallel β -sheets packed at approximately 90° to one another create an aromatic patch on the surface of the SH3 domain for binding to proline-rich ligands. In analogy to intermolecular SH3–ligand complexes^{13–15}, the internal KLPPTP ligand of Itk adopts a left-handed PPII helix where the Leu-Pro and Thr-Pro dipeptide moieties contact the SH3 binding pocket directly (Fig. 2*a*). NOE contacts between KLPPTP and the SH3 binding pocket of

ItkSH3-KPL are readily identified by comparison with spectra of ItkSH3, and indicate that KLPPTP binds in a class I orientation (Fig. 2*b*). Additionally, the critical prolines of KLPPTP are indispensable for SH3 binding; mutation to alanine prohibits formation of the folded structure (Fig. 2*c*). No long- or medium-range NOEs are observed for the loop residues connecting the SH3 domain and the KLPPTP ligand. This intervening Tec loop, which is present in all Tec family members, is not well ordered in this fragment of Itk but tends to occupy the quadrant between the n-Src loop and the carboxy terminus of the SH3 domain (Fig. 1*c*). The structure of this intramolecular complex suggests that the proline-rich region of Itk may regulate binding of the SH3 domain to its cellular targets.

To assess the relevance of this intramolecular complex *in vitro*, we generated a panel of glutathione *S*-transferase (GST)–Itk fusion proteins (Fig. 3*a*). Several binding partners of the Itk SH3 domain have been identified previously¹⁶, and the ability of fusion proteins containing the Itk SH3 domain to bind these ligands served as a measure of the accessibility of the SH3 domain. Consistent with the relatively low-affinity association observed by NMR (Fig. 2*c*), only a slight reduction in the accessibility of the Itk SH3 domain was observed for the ItkSH3–KPL fusion protein (data not shown). In contrast, addition of the TH domain dramatically reduced binding of a large number of ³⁵S-labelled proteins to the Itk SH3 domain (Fig. 3*b*). These domains were relatively ineffective at inhibiting the binding of a high-affinity ligand, Sam68, to the Itk SH3 domain (Fig. 3*c*; compare SH3, TH3 and TEC3). This is consistent with a model in which the Itk proline-rich region can compete with low-to-intermediate-affinity cellular ligands for

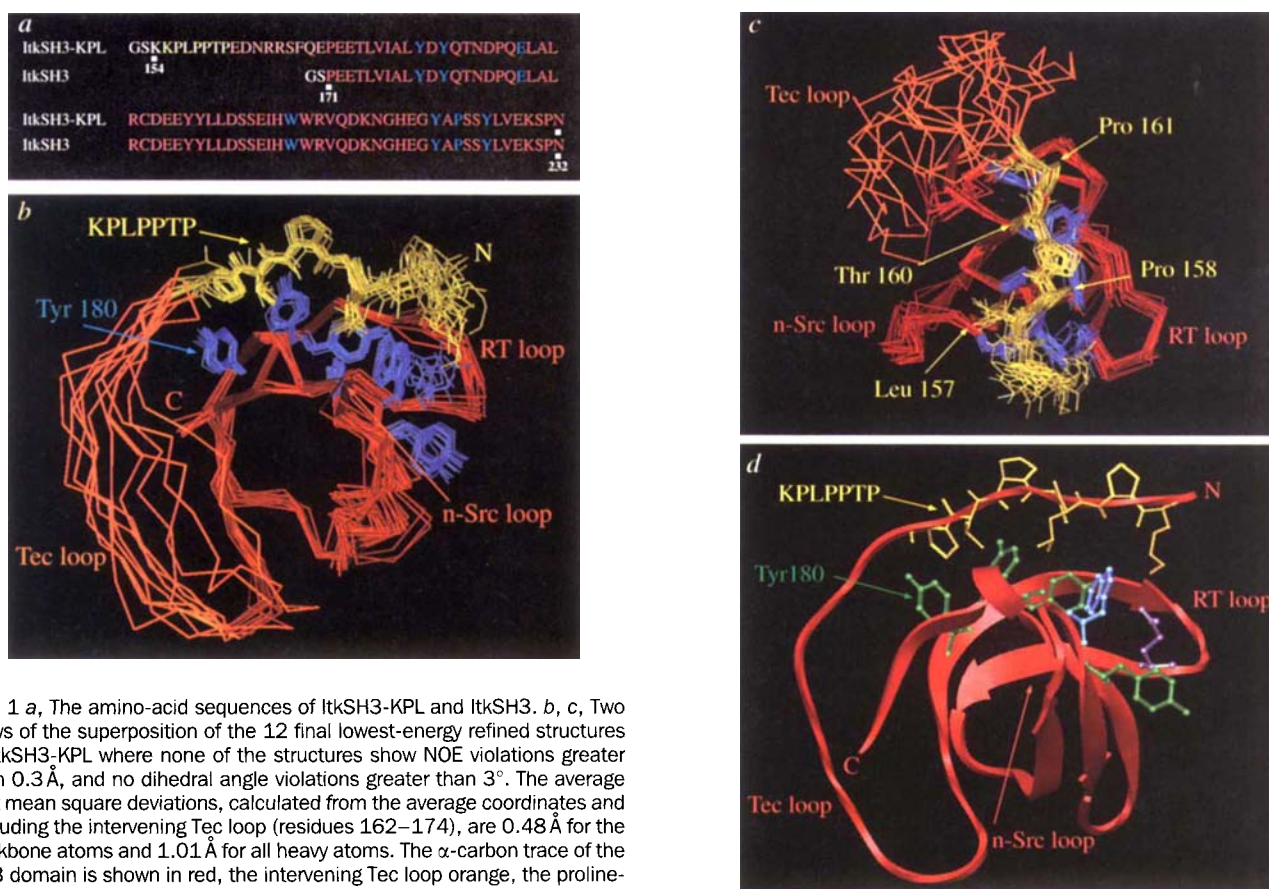


FIG. 1 *a*, The amino-acid sequences of ItkSH3-KPL and ItkSH3. *b*, *c*, Two views of the superposition of the 12 final lowest-energy refined structures of ItkSH3-KPL where none of the structures show NOE violations greater than 0.3 Å, and no dihedral angle violations greater than 3°. The average root mean square deviations, calculated from the average coordinates and excluding the intervening Tec loop (residues 162–174), are 0.48 Å for the backbone atoms and 1.01 Å for all heavy atoms. The α -carbon trace of the SH3 domain is shown in red, the intervening Tec loop orange, the proline-rich sequence (KLPPTP) yellow, and the conserved aromatic side chains in the SH3 binding pocket are blue. *d*, Ribbon trace (red) of the refined average intramolecular complex depicting important residues of the Itk SH3 binding pocket and the internal Itk ligand, KLPPTP (yellow). Tyrosine residues are

shown in green, Trp 208 on the n-Src loop light blue, and Glu 189, the conserved acidic site located on the RT loop, is purple. The site corresponding to the autophosphorylation site of Btk²³ is Tyr 180 (green).

binding to the Itk SH3 domain.

The binding properties of fusion proteins containing the Itk SH2 domain suggest that this domain can stabilize the intramolecular interaction between the SH3 domain and proline-rich region. Specifically, the ability of the Itk SH3 domain to bind Sam68 is strongly inhibited in fusion proteins containing both the Itk SH2 domain and the amino-terminal region of Itk (Fig. 3c; compare TH3 and TEC3 with TH2 and TEC2). However, the Itk SH2 domain does not suppress the ability of the Itk SH3 domain to bind Sam68 in the absence of the N-terminal region of Itk (Fig. 3c; compare SH3 and SH32). It is likely that the Itk SH2 domain does not directly inhibit binding of Sam68 to the Itk SH3 domain, but may interact with the N-terminal region of Itk to stabilize the described inhibitory interaction. Furthermore, the involvement of the Itk proline-rich region in this inhibitory interaction prevents its association with a ligand containing an SH3 domain. Grb-2, which is capable of binding the proline-rich region of Itk through its N-terminal SH3 domain (S.C.B. and L.J.B., unpublished results), does not bind to the Itk proline-rich region in the context of the TEC2 fusion protein (Fig. 3d; compare TEC and TEC2). Thus the SH3 domain and proline-rich region of Itk are capable of

interacting intramolecularly to inhibit the binding of cellular ligands such as Sam68 and Grb-2, respectively.

The mutation of critical residues opens this structure, allowing both the proline-rich region and the binding pocket of the SH3 domain to interact with other proteins. For example, binding of the Itk SH3 domain to Sam68 is restored upon mutation of the proline-rich region (KPLPPTP to KPLAATP; TEC2:PR* in Fig. 3d), while binding of the proline-rich region to Grb-2 is recovered upon mutation of the critical Trp 208 residue in the ligand-binding pocket of the SH3 domain (TEC2:3* in Fig. 3d). Furthermore, mutation of the TH domain reveals both Sam68 and Grb-2 binding activities, indicating that the TH domain, like the SH2 domain, is involved in stabilizing the closed intramolecular complex (TEC2:TH* in Fig. 3d). The NMR structure of the ItkSH3-KPL fragment demonstrates clearly that Itk can fold back on itself in the region of the Tec loop (Fig. 1d), so it is possible that the TH domain contacts the SH2 domain and further stabilizes the observed interaction between the SH3 domain and the proline-rich region. Thus mutational analysis of the Itk SH3, TH and proline-rich domains is consistent with an intramolecular SH3-proline complex that excludes potential cellular ligands.

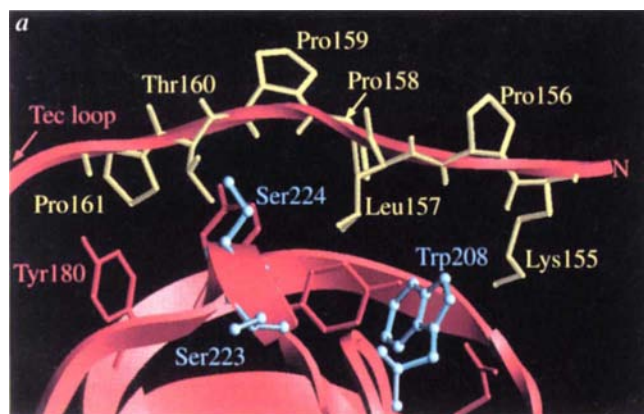
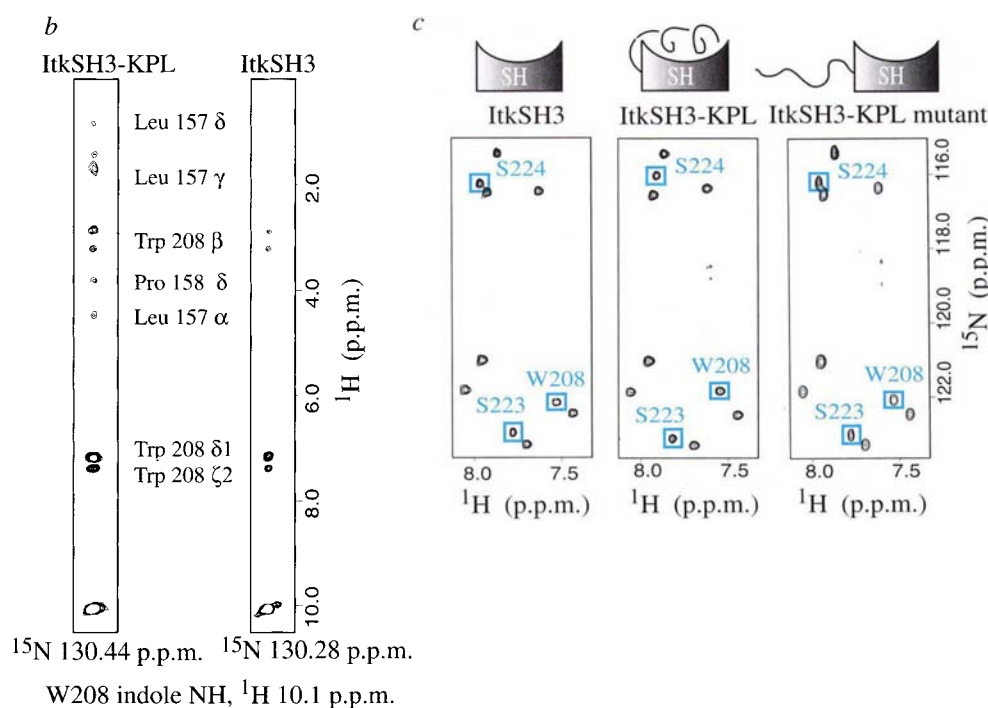


FIG. 2 a, The interface of the intramolecular SH3-ligand complex of Itk. The KPLPPTP sequence is shown in yellow and Lys 155 to Pro 161 are labelled. Residues Ser 223, Ser 224 and Trp 208 are shown in light blue and correspond to the chemical shift perturbation discussed in c. b, Comparison of three-dimensional NOESY $^1\text{H}/^{15}\text{N}$ -HMQC spectra for ItkSH3-KPL and ItkSH3. The ^{15}N slice corresponding to the chemical shift of the W208 indole nitrogen is shown. Additional NOEs are observed for ItkSH3-KPL and correspond to protons of the Leu 157-Pro 158 dipeptide of the KPLPPTP sequence in the proline-rich region of Itk. α , β , γ and δ refer to the position of the side-chain protons of Leu 157 and Pro 158. c, Chemical shift analysis indicates that the association of the KPLPPTP sequence and the adjacent SH3 domain is dependent on PPII helix formation. Portions of the ^1H - ^{15}N HSQC spectra with assignments of selected peaks are shown for recombinant ItkSH3, ItkSH3-KPL and a mutant in which the critical prolines in the KPLPPTP sequence of ItkSH3-KPL have been mutated to alanine (KPLAPTA). Several amide resonance frequencies (^1H and ^{15}N) differ between the ItkSH3 and ItkSH3-KPL proteins. The chemical shift perturbations shown (W208, S223 and S224) are localized to the SH3 binding pocket, as shown in a. The magnitude of each chemical shift perturbation is small, indicating a low-affinity association and entirely consistent with the binding behaviour of the ItkSH3-KPL fusion protein. Abolishing PPII helix formation (ItkSH3-KPLmutant) results in resonance frequencies identical to those of ItkSH3, suggesting that the proline-rich sequence no longer binds to the SH3 binding pocket.



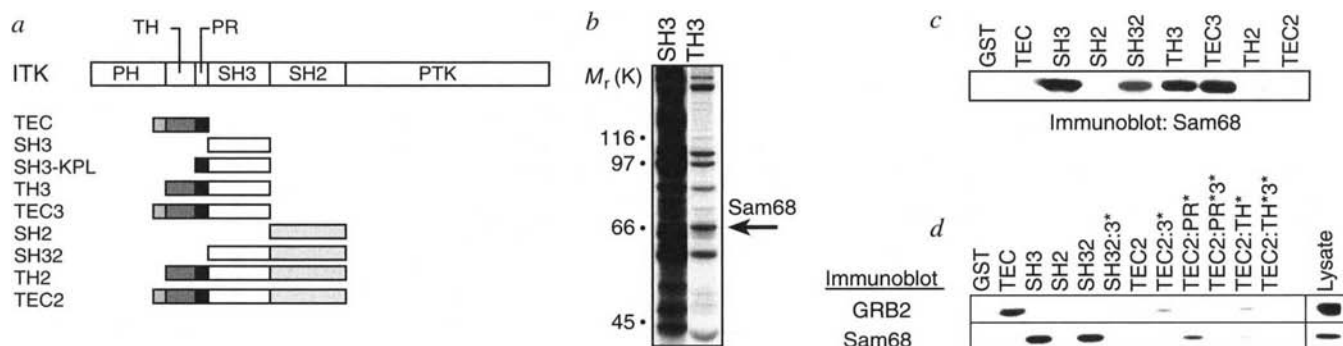


FIG. 3 The ability of the SH3 domain and proline-rich region of Itk to bind their respective cellular ligands, Sam68 and Grb-2, is inhibited by an intramolecular interaction. *a*, Domain structure of Itk and the GST-Itk fusion proteins. PH refers to the pleckstrin homology domain and TH (Tec homology) designates the region homologous to the GAP1 family of Ras GTPases. The proline-rich domain is designated PR. *b*, Binding of ^{35}S -Met-

labelled Jurkat proteins to Itk fusion proteins. *c*, Itk fusion proteins (2 μg) were incubated with Jurkat lysates. Bound proteins were immunoblotted for Sam68. *d*, Itk fusion proteins (2 μg) were incubated with lysates of mock-stimulated Jurkat cells. The accessibility of the SH3 domain and the proline-rich region was monitored by immunoblotting for Sam68 and Grb-2, respectively.

Although the SH2 domain stabilizes the intramolecular complex, the phosphotyrosine (pTyr)-binding pocket of the Itk SH2 domain appears accessible in this complex, unlike the binding pocket of the Itk SH3 domain. This is most clearly illustrated by the similar pattern of pTyr-containing proteins bound by the SH2 and TEC2 fusion proteins (Fig. 4*a*). However, in the absence of the N-terminal region, extensive synergistic binding is observed between the SH3 and SH2 domains (Fig. 4*a*; compare SH32 and SH32:3*). Thus the N-terminal region of Itk suppresses the ability of the Itk SH3 domain to synergize with the Itk SH2 domain in binding tyrosine phosphoproteins. It is likely that the intramolecular Itk proline-rich region competes with cellular ligands that contain binding sites for both the SH2 and SH3 domains of Itk. We propose that there exists an equilibrium between an intramolecular, 'closed' complex and intermolecular, 'open' complexes in which the Itk SH3 and SH2 domains are occupied by bidentate cellular ligands. Furthermore, we suggest that the balance between 'closed' and 'open' forms could be manipulated by altering the relative abundance of such ligands.

We tested this hypothesis by exposing the fusion proteins to lysates containing varying levels of tyrosine phosphorylated proteins. The ability of cellular Grb-2 to bind the TEC2 fusion protein was used to assess the accessibility of the Itk proline-rich region, and was thus an indirect measure of whether the resulting complexes were 'open' or 'closed'. The TEC2 fusion protein does not bind Grb-2 from unstimulated cells, but TEC2 binds Grb-2 weakly in TCR-stimulated lysates and very strongly in pervanadate-treated lysates (Fig. 4*b*). As the amount of Grb-2 bound to the proline-rich region of the TEC2 fusion protein increases with the level of tyrosine phosphorylation present in the lysates, it seems likely that ligands binding to the Itk SH3 and SH2 domains outcompete the intramolecular interaction between the Itk SH3 domain and proline-rich region. This equilibrium between 'open' and 'closed' states may exist *in vivo*, where the balance between 'open' and 'closed' states could be further modulated by colocalization of Itk with putative ligands, or by potential interactions of the PH and tyrosine-kinase domains present in full-length Itk (Fig. 5*a, b*).

We have identified an intramolecular SH3-ligand interaction that regulates the binding of the Itk SH3 domain and proline-rich regions to their respective targets *in vitro*. The pTyr-binding pocket of the SH2 domain appears accessible in this complex, potentially enabling ligands of the Itk SH3 and SH2 domains to sever the observed intramolecular interaction, revealing a proline-rich SH3-binding site. In a similar fashion, it has been demonstrated that the accessibility of an SH3 domain-binding site in the

adaptor protein Crk-II is modulated by binding of a pTyr-containing ligand to the Crk-II SH2 domain²². Furthermore, it is possible that direct phosphorylation of the Itk proline-rich region or SH3 domain may disrupt the observed intramolecular complex, exposing binding sites for interactions with other proteins. For example, Tyr 180 within the SH3 domain of Itk corresponds to the autophosphorylation site of the Tec family member Btk²³, and

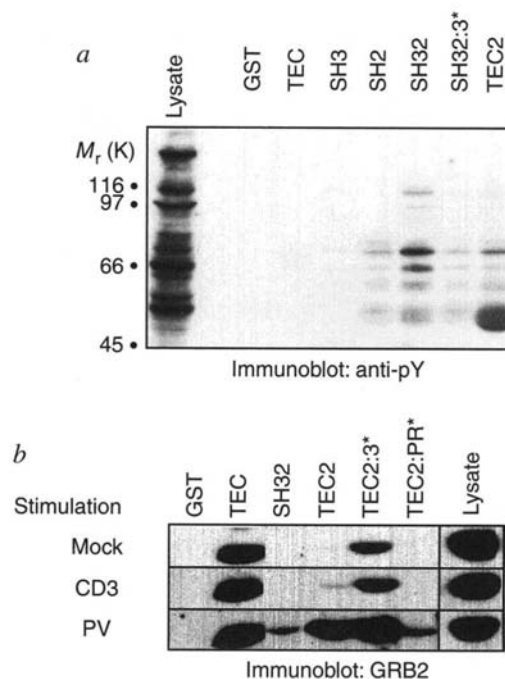


FIG. 4 The phosphotyrosine-binding pocket of the Itk SH3 domain remains accessible in the intramolecular complex and may be involved in the disruption of this complex. *a*, Itk fusion proteins (2 μg) were incubated with lysates of TCR-stimulated Jurkat cells. Bound proteins were immunoblotted with the anti-phosphotyrosine monoclonal antibody 4G.10. The large band at approximately 55K in the TEC2 lanes is the result of a crossreaction between the secondary antibody and the TEC2 fusion proteins. *b*, Itk fusion proteins (2 μg) were incubated with lysates of mock-, TCR- or pervanadate (PV)-stimulated Jurkat cells. Bound proteins were immunoblotted with anti-Grb-2.

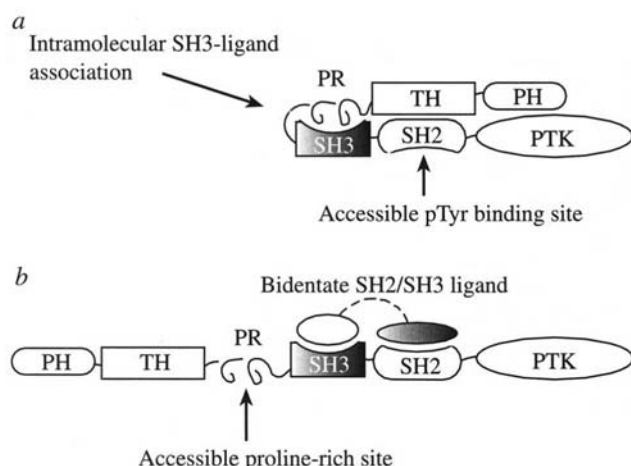


FIG. 5 *a*, Model of the observed intramolecular interaction, showing the observed interaction between the Itk proline-rich region and SH3 domain as well as an interaction between the Itk TH and SH2 domains. *b*, Model of the opening of the intramolecular complex by interaction with a bidentate ligand for the Itk SH3 and SH2 domains.

points directly at acidic residues within the Tec loop (Fig. 1*d*). In addition, the proline-rich region of Itk contains predicted MAP kinase²⁴, casein kinase II²⁵ and protein kinase C²⁶ phosphorylation sites. An analogy for this intramolecular model is found in the Src family kinases, in which phosphorylation of a C-terminal tyrosine residue is thought to regulate kinase activity and substrate binding^{9,10} through an intramolecular association with the SH2 domain. Similar intramolecular regulatory interactions have also been identified in the adaptor protein Crk²⁷. It is provocative that Tec family members do not contain the regulatory phosphorylation site present in Src family kinases, but may be regulated instead through the described intramolecular interaction between the SH3 domain and proline-rich region. We suggest that this may be a general mechanism by which the activity of Tec family kinases is controlled. □

Methods

Preparation of fusion proteins. Inserts encoding the desired domains of Itk were generated by PCR and were subcloned into the *Bam*HI site of pGEX-2T. The TEC, TH, SH3-KPL, SH3 and SH2 5' primers encode Itk products beginning at amino acids 97, 109, 154, 171 and 231, respectively. The residue numbering system used throughout the text is that of full-length Tsk/Itk¹. Point mutations were introduced by a combination of PCR-directed mutagenesis and site-directed mutagenesis. The 3* mutation inactivates the proline-binding pocket of the SH3 domain by converting Trp 208 to Lys. The TH* mutation converts two highly conserved Cys residues in the TH domain (Cys 132 and Cys 133) to Arg residues. The PR* mutation converts Pro 158 and Pro 159 to Ala residues. All fusion proteins were produced in *E. coli* DH5 α ¹⁶, with the exception of TEC2 and variants thereof. To obtain suitable levels of undegraded protein, the TEC2 fusion proteins were grown in the *E. coli* strain BL21(DE3) in the presence of pUBS1560, a kanamycin-resistant vector that encodes tRNAs for the rare Arg codons AGA and AGG²⁸.

Collection of NMR spectra and determination of structure. Uniformly labelled samples of ¹⁵N/¹³C ItkSH3 and ItkSH3-KPL were expressed using the pGEX-2T expression vector and purified^{13–15}. The NMR samples contained ~2.5 mM protein in a D₂O or 90% H₂O/10% D₂O buffer containing 50 mM potassium phosphate, pH 6.0, 150 mM KCl, 0.1 mM DTT, 0.1 mM EDTA and 0.02% NaN₃. Spectra were recorded on a Bruker DMX500 spectrometer at 300 K. Protein resonances were assigned by a series of two- and three-dimensional NMR experiments^{13–15}. Distance (NOE) and dihedral (ϕ and χ_1) restraints derived from the NMR experiments were used to calculate structures with the program X-PLOR²⁹. To calculate the family of structures, a total of 672 NOE restraints, 22 of which were identified between the KPLPTTP sequence and the Itk SH3 binding pocket, and 55 dihedral restraints were used. Hydrogen-bond distances (based on slowly exchanging amide protons) and a salt bridge restraint (E189–K155, based on previous SH3-ligand structures^{13,14}) were included in the final stage of calculation. The average structure was calculated

from the coordinates of the 12 final structures and refined with 500 steps of steepest-descent energy minimization with the simulated annealing parameters. The van der Waals energy of the refined average structure is $-47.6 \text{ kcal mol}^{-1}$, as calculated by CHARMM-19 parameters in X-PLOR²⁹. The structures are drawn with the program Ribbons 2.2 (ref. 30).

Preparation of Jurkat lysates and binding assays. Jurkat cells were labelled with ³⁵S-Met and lysed¹⁶. Tyrosine phosphorylated lysates were prepared from Jurkat cells mock- or OKT3-stimulated for 2 min (ref. 16), or from Jurkat cells stimulated for 10 min at 37 °C with 500 μM pervanadate. All binding assays were performed and analysed as previously described¹⁶.

Immunoblots. All immunoblots were performed as described previously¹⁶. Antibodies to Sam68 and Grb-2 were from Santa Cruz Biotechnology (sc-333 and sc-255).

Received 18 September; accepted 21 November 1996.

- Heyeck, S. D. & Berg, L. J. *Proc. Natl. Acad. Sci. USA* **90**, 669–673 (1993).
- Siliciano, J. D., Morrow, T. A. & Desiderio, S. V. *Proc. Natl. Acad. Sci. USA* **89**, 11194–11198 (1992).
- Tanaka, N., Asao, H., Ohtani, K., Nakamura, M. & Sugmua, K. *FEBS Lett.* **324**, 1–5 (1993).
- Mano, H. et al. *Oncogene* **8**, 417–424 (1993).
- Tsukada, S. et al. *Cell* **72**, 279–290 (1993).
- Vetrie, D. et al. *Nature* **361**, 226–232 (1993).
- Tamagnone, L. et al. *Oncogene* **9**, 3683–3688 (1994).
- Liao, X. C. & Littman, D. R. *Immunity* **3**, 757–769 (1995).
- Courtneidge, S. A. *EMBO J.* **4**, 1471–1477 (1985).
- Cooper, J. A. & Howell, B. *Cell* **73**, 1051–1054 (1993).
- Chen, J. K., Lane, W. S., Brauer, A. W., Tanaka, A. & Schreiber, S. L. *J. Am. Chem. Soc.* **115**, 12591–12592 (1993).
- Rickles, R. J. et al. *Proc. Natl. Acad. Sci. USA* **92**, 10909–10913 (1995).
- Yu, H. et al. *Cell* **76**, 933–945 (1994).
- Feng, S., Chen, J. K., Yu, H., Simon, J. A. & Schreiber, S. L. *Science* **266**, 1241–1247 (1994).
- Feng, S., Kasahara, C., Rickles, R. J. & Schreiber, S. L. *Proc. Natl. Acad. Sci. USA* **92**, 12408–12415 (1995).
- Bunnell, S. C. et al. *J. Biol. Chem.* **271**, 25646–25656 (1996).
- Cheng, G., Ye, Z. & Baltimore, D. *Proc. Natl. Acad. Sci. USA* **91**, 8152–8155 (1994).
- Ikura, M. et al. *Science* **256**, 632–638 (1992).
- Marion, D. et al. *Biochemistry* **28**, 6150–6156 (1989).
- Zuiderweg, E. R. P. & Fesik, S. W. *Biochemistry* **28**, 2387–2391 (1989).
- Chen, J. K. & Schreiber, S. L. *Bioorg. Med. Chem. Lett.* **4**, 1755–1760 (1994).
- Anafi, M., Rosen, M. K., Gish, G. D., Kay, L. E. & Pawson, T. *J. Biol. Chem.* **271**, 21365–21374 (1996).
- Park, H. et al. *Immunity* **4**, 515–525 (1996).
- Crews, C. M., Alessandrini, A. & Erikson, R. L. *Cell Growth Differ.* **3**, 135–142 (1992).
- Pinna, L. A. *Biochim. Biophys. Acta* **1054**, 267–284 (1990).
- Woodgett, J. R., Gould, K. L. & Hunter, T. *Eur. J. Biochem.* **161**, 177–184 (1986).
- Rosen, M. K. et al. *Nature* **374**, 477–479 (1995).
- Brinkmann, U., Mattes, R. E. & Buckel, P. *Gene* **85**, 109–114 (1989).
- Brunker, A. T. *X-PLOR 3.1* (Yale Univ. Press, New Haven, CT, 1992).
- Carson, M. J. *Appl. Crystallogr.* **24**, 958–961 (1991).

ACKNOWLEDGEMENTS. A.H.A. and S.C.B. contributed equally to this work. We thank B. Drucker for the anti-phosphotyrosine monoclonal antibody 4G.10. A.H.A. is supported by an NIH postdoctoral fellowship and the Bunting Institute of Radcliffe College. S.L.S. is an investigator at the Howard Hughes Medical Institute. This work was supported by the NIH General Medical Sciences (S.L.S.), the NIH NIAID (L.J.B.) and the Council for Tobacco Research Inc. (L.J.B.).

CORRESPONDENCE and requests for materials should be addressed to S.L.S. (e-mail: sls@slsiris.harvard.edu).

CORRECTION

Identification of the homologous beige and Chediak–Higashi syndrome genes

Maria D. F. S. Barbosa, Quan A. Nguyen, Velizar T. Tchernev, Jennifer A. Ashley, John C. Detter, Susan M. Blaydes, Stephen J. Brandt, Dipti Chotai, Charles Hodgman, Roberto C. E. Solari, Michael Lovett & Stephen F. Kingsmore

Nature **382**, 262–265 (1996)

THE Genbank accession number of the mouse beige gene (*Lyst*) was incorrect. It is L77884. The Genbank accession number of the human Chediak–Higashi syndrome gene (*LYST*) is L77889.

Guide to authors of contributions to *Nature*

Nature is a weekly international journal covering all the sciences. It is intended for an interdisciplinary readership, so all manuscripts should be written clearly and simply. Contributors should pay particular attention to the fact that English is not the first language of many readers. Space is limited and competition severe, so brevity is highly valued. The length limits below are specified in terms of the number of pages taken up in *Nature*. An uninterrupted page of text contains about 1,300 words, so a typical Letter to *Nature*, occupying 2½ pages, can contain about 1,500 words of text and four modest display items (figures and/or tables) with brief legends. There are no formal limits on the number of display items, but extra display material must be accompanied by commensurate reductions in the length of the text. Length limits can, however, be varied upwards or downwards at the editors' discretion.

Nature publishes the following types of contribution:

Review Articles survey recent developments in a topical area of scientific research. They should not generally occupy more than 6 pages of *Nature*.

Progress articles are relatively informal reviews of particularly topical and fast-moving fields. They do not exceed 4 pages of *Nature*.

Contributors wishing to submit a Review Article or Progress should first send a one-page synopsis to the Reviews Coordinator, from whom guidelines are available on request.

Articles are reports of original scientific research whose conclusions represent a substantial advance in understanding of an important problem and are of broad general interest. They do not generally exceed 5 pages of *Nature* and have no more than 50 references. They have a 'heading' of up to 80 words, often rewritten by *Nature's* editors, which advertises the content of the paper in general terms and should not contain numbers, abbreviations or measurements unless central to the message of the paper. The introduction and summary should be contained in the first three or four paragraphs of the main text. Layout of methods, figures and legends is summarized below.

Letters to *Nature* are short reports of original research focused on an outstanding finding whose implications are of interest to nonspecialists. They should not generally exceed 2½ pages of *Nature*. They have a bold first paragraph of not more than 180 words summarizing the background, rationale, conclusions and main implications of the work, aimed particularly at nonspecialist readers. Letters should have no more than 30 references. Layout of methods, figures and legends is summarized below.

Preparation of Articles and Letters

All manuscripts should be written in English, typed double-spaced on one side of the paper only. Manuscripts submitted in electronic form cannot be considered.

Manuscripts should be accompanied by a brief covering letter from the corresponding author, containing a full postal address, telephone and fax numbers, and e-mail address.

Five copies of manuscripts and original figures are needed, together with two copies of the covering letter. Five copies of manuscripts by any of the authors that have been submitted or are in press elsewhere should also be included, clearly marked as such; failure to do so may prejudice consideration of the manuscript.

Unless otherwise instructed by the editor handling a manuscript, when sending revised or resubmitted manuscripts five copies are required, each accompanied by a copy of the authors' response to referees' and editors' comments on the earlier version.

Titles should not exceed three lines of 31 characters each (including spaces), and should avoid numerical values, acronyms, abbreviations and punctuation wherever possible.

References should be numbered sequentially as they appear in the text, tables and figure legends. Only papers that have been published or submitted to a journal should be in the numbered list: manuscripts in preparation should be mentioned in the text with a list of authors. Reference lists should not contain textual material, grant details or acknowledgements.

Acknowledgements must be brief and appear after the reference list. *Nature* does not publish grant or contribution numbers.

Figures should not be larger than 22 by 28 cm unless unavoidable. They should not contain more than one panel unless the assembled parts are logically connected. Each copy should be marked with the corresponding author's name and, when known, the manuscript reference number. One photocopy of the original figures should be provided. All original figures are returned when *Nature* cannot offer to

publish a manuscript, but one copy of the manuscript and a photocopy of the figures are kept in confidential files for three months and then destroyed. Detailed instructions for figure preparation are available from the production department, including details of acceptable digital formats.

Colour figures are welcome but a contribution towards the cost of reproduction is requested. Inability to pay this charge will not prevent *Nature* publishing figures where colour is essential.

Tables. Symbols and abbreviations in a table should be defined immediately below the table, followed by essential descriptive material, all in double-spaced text. Tables should each be presented on a separate sheet of paper.

Methods. Brief details of materials and methods can be given in the text. Otherwise, they should be described at the end of the text in a section entitled "Methods", subdivided by bold headings referring to the methods used. This section should not exceed 800 words and ideally should be shorter. Additional material can be deposited on *Nature's* Web site. In contributions with methods sections, figure legends should not exceed 150 words each, and should be confined to a description of the figure and the symbols used.

Materials. As a condition of publication, authors are required to make materials and methods used freely available to academic researchers for their own use. Supporting datasets must be made available at the time of publication from the authors directly, by posting on *Nature's* Web site, by deposition in the appropriate public database or by distribution on the Internet. In the case of X-ray crystallographic coordinates only, public access may be delayed for up to one year after publication.

Articles and Letters should both contain a statement at the end of the text: "Correspondence and requests for materials to xxxxx", with an e-mail address if desired, followed by the accession numbers or Web site address for supporting datasets. The designated author need not necessarily be the senior author, but *Nature* will expect the named author to handle matters arising from published contributions.

Manuscripts can be submitted to the Editor at *Nature*, Porter's South, Crinan St, London N1 9XW or at *Nature*, 968 National Press Building, Washington, DC 20045-1938. Proofs should be returned by express mail to London. Fax numbers and e-mail addresses for all editorial offices are on *Nature's* masthead.

Other contributions

Nature publishes informal contributed material as follows:

Correspondence contributions are short comments on topical issues, anecdotal material or reactions of readers to nontechnical material published in *Nature* (mainly News, Briefings, Opinion, Book Review or Commentary articles). All are unsolicited; submissions by e-mail to corres@nature.com are preferred (accents on names and places should be clearly indicated).

Commentary articles are informal, editorial comments on topical issues of public interest that have some bearing on scientific research. They are often controversial. Unsolicited manuscripts are considered, but enquiries can be made to the Commentary Editor by e-mail (nature@nature.com) or fax as a one-page synopsis or informal letter before formal submission.

News and Views articles inform nonspecialist readers about new scientific advances, either as reported in published papers or in the form of meeting reports. Most are commissioned, but proposals can be made to the News and Views Editor in advance of publication of the paper or of the meeting concerned. Authors are not allowed to discuss their own work or work from elsewhere in their institution. Detailed guidelines are available on request.

Scientific Correspondence is a predominantly peer-reviewed section of the journal specializing in very brief reports of general interest and discussion of topical issues. Priority is given to contributions of fewer than 500 words. Detailed guidelines are available on request. All contributions are unsolicited, and should be submitted by mail, not e-mail. Copies of contributions submitted by fax should be sent by mail with original figures and a cover letter noting that the manuscript has been faxed.

Book Reviews are all commissioned by the Book Review Editor. Unsolicited contributions are not considered.

Nature's editors will request final copy for Commentary, News and Views, Scientific Correspondence and Book Reviews on disk or via e-mail, as appropriate.

A more detailed Guide to Authors is available on *Nature's* WWW sites at <http://www.nature.com> or <http://www.america.nature.com>